



PCT

**WORLD INTELLECTUAL PROPERTY ORGANIZATION**  
**International Bureau**

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :  C12N 15/12, 15/31, 15/62, 1/19, C07K 14/47, 14/395, C12Q 1/02, 1/68, G01N 33/50		A1	(11) International Publication Number:  WO 99/18211
(21) International Application Number:  PCT/US98/21168		(43) International Publication Date:  15 April 1999 (15.04.99)	
(22) International Filing Date: 7 October 1998 (07.10.98)			
(30) Priority Data:  08/946,298 7 October 1997 (07.10.97) US		tine, A. [US/US]; Teatown Lake Reservation, Spring Valley Road, Ossining, NY 10562 (US). MURPHY, Andrew, J., M. [US/US]; 17 Windsor Place, Montclair, NJ 07043 (US). XU, Jun [CN/US]; 102 Morningside Drive, Ossining, NY 10562 (US). BENEGAL, Anupama, N. [IN/US]; 437 The Fenway, River Edge, NJ 07661 (US).	
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application  US 08/946,298 (CON) Filed on 7 October 1997 (07.10.97)		(74) Agents: DECONTI, Giulio, A., Jr. et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).	
(71) Applicant (for all designated States except US): CADUS PHARMACEUTICAL CORPORATION [US/US]; 777 Old Saw Mill River Road, Tarrytown, NY 10591 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(72) Inventors; and (75) Inventors/Applicants (for US only): BROACH, James, R. [US/US]; 6 Andrews Lane, Princeton, NJ 08540 (US). MANFREDI, John, P. [US/US]; Teatown Lake Reservation, Spring Valley Road, Ossining, NY 10562 (US). PAUL, Jeremy, I. [US/US]; 89 Depot Place, South Nyack, NY 10690 (US). TRUEHART, Joshua [US/US]; 212 South Broadway, South Nyack, NY 10960 (US). KLEIN, Chris-		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	

(54) Title: YEAST CELLS EXPRESSING MODIFIED G PROTEINS AND METHODS OF USE THEREFOR

Sequence alignments of N-terminal regions of Ga subunits and N-terminal sequences of  $\text{GPA}_4\text{-Ga}$  hybrid proteins.

#### A. Alignment of GPa1 with Gα Subunits

```

GPA1
KGC.TVSTQTIGDESDPFLOMIRARNDVIEQSIOLIKRQPKRMTKILLLLGAGTSGKSTTVLHQIILQQ...
Gc8
KGCLGTS...RTYDGRHKKQGPZANHGCKQGKQDGVYRATHPILLLGAGTSGKSTTVQD-BILHV...
Gc12
KGC.TVS.....AEDKAALVERSIMIDENLIEDGKAAAEIVKLLLLLGAGTSGKSTTVQKMKIHE...
Gc13
KGC.TVS.....AEDKAALVERSIMIDENLIEDGKAAAEIVKLLLLLGAGTSGKSTTVQKMKIHE...
Gc16
MARSLSLWRCCPNCLTDEKAAARVDOZTRILLEORKODRGKILKLLLLLGPGRSGKSTTVQMRISG...

```

### B. $GPA_{41}$ -Ga Hybrids

GPA<sub>41</sub> - Gα8  
KCC: TVSTVTTGDESDFYLQNKGRANDVIEQSLOLLKORDKEKREKLILLLGCGSGKSTIVQMRILIV.  
GPA<sub>41</sub> - Gα12  
KCC: TVSTVTTGDESDFYLQNKGRANDVIEQSLOLKGDKDNEVILILLLGCGSGKSTIVQMRILIV.  
GPA<sub>41</sub> - Gα13  
KCC: TVSTVTTGDESDFYLQNKGRANDVIEQSLOLKGDKDNEVILILLLGCGSGKSTIVQMRILIV.  
GPA<sub>41</sub> - Gα16  
KCC: TVSTVTTGDESDFYLQNKGRANDVIEQSLOLKGDKDNEVILILLLGCGSGKSTIVQMRILIVG

Applicants: Mary Cismowski, et al.  
Serial No.: 09/709,103  
Filed: November 8, 2000  
**Exhibit 19**

## (57) Abstract

The present invention pertains to novel yeast cells which are useful for the expression of heterologous G protein coupled receptors. The yeast cells of the present invention can be used in screening assays which can be used to screen for modulators of G protein coupled receptors. Specifically, the invention provides novel yeast cells which express a heterologous G protein coupled receptor and mutant and/or chimeric G protein subunit molecules which serve to functionally integrate the heterologous into the pheromone signaling pathway of the yeast cell. The invention also provides for the expression of heterologous G protein coupled receptors which are functionally integrated into the yeast cell membrane using a yeast  $\alpha$  factor leader sequence. Drug discovery assays using the subject yeast cells are also provided.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## YEAST CELLS EXPRESSING MODIFIED G PROTEINS AND METHODS OF USE THEREFOR

### 5 Cross Reference to Related Applications

This application is a continuation-in-part of USSN 08/946,298, filed on October 7, 1997, now pending, which is a continuation-in-part of USSN 08/689,172, filed on August 6, 1996 now pending, which is a continuation-in-part of USSN 08/582,333, filed January 17, 1996, now pending, which is a continuation-in-part of 10 USSN 08/463,181, filed June 5, 1995, now pending, which is a continuation-in-part of USSN 08/322,137, filed October 13, 1994, now pending, which is a continuation-in-part of USSN 08/309,313, filed September 20, 1994, now abandoned, which is a continuation-in-part of USSN 08/190,328, filed January 31, 1994, now abandoned, which is a continuation-in-part of USSN 08/041,431, filed March 31, 1993, now 15 abandoned, the specifications of which are hereby incorporated by reference.

### Background Of The Invention

Cell surface receptors are an important class of proteins involved in cellular functioning because they are the primary mediators of cell to cell communication. G protein coupled receptors (GPCRs) are an important category of cell surface receptors. The medical importance of these receptors is evidenced by the fact that more than 60% of all commercially available prescription drugs work by interacting with known GPCRs.

In their resting state, the G proteins, which consist of alpha ( $\alpha$ ), beta ( $\beta$ ) 25 and gamma ( $\gamma$ ) subunits, are complexed with the nucleotide guanosine diphosphate (GDP) and are in contact with the receptors to which they are coupled. When a hormone or other first messenger binds to receptor, the receptor changes conformation and this alters its interaction with the G protein. This spurs the  $\alpha$  subunit to release GDP, and the more abundant nucleotide guanosine triphosphate (GTP) replaces it, activating the G 30 protein. The G protein then dissociates to separate the  $\alpha$  subunit from the still complexed beta and gamma subunits. Either the  $G\alpha$  subunit, or the  $G\beta\gamma$  complex, depending on the pathway, interacts with an effector. The effector (which is often an enzyme) in turn converts an inactive precursor molecule into an active "second messenger," which may diffuse through the cytoplasm, triggering a metabolic cascade. 35 After a few seconds, the  $G\alpha$  converts the GTP to GDP, thereby inactivating itself. The inactivated  $G\alpha$  may then reassociate with the  $G\beta\gamma$  complex.

- 2 -

Hundreds, if not thousands, of receptors convey messages through heterotrimeric G proteins, of which at least 17 distinct forms have been isolated. Most G protein-coupled receptors are comprised of a single protein chain that is threaded through the plasma membrane seven times. Such receptors are often referred to as seven-transmembrane receptors (STRs). More than a hundred different GPCRs have been found, including many distinct receptors that bind the same ligand, and there are likely many more GPCRs awaiting discovery. The development of new drug discovery assays to identify novel modulators of GPCRs would be of tremendous benefit.

In recent years drug discovery has been advanced by expression of heterologous receptors in living cells. However, due to the complexity of GPCRs the search for modulators of these receptors have presented particular challenges. For example, there is variability in the sequences of G protein subunits and this variability can influence the efficiency of receptor coupling to subunits. The highest variability has been seen in the  $\alpha$  subunit, but several different  $\beta$  and  $\gamma$  structures have also been reported. There are, additionally, several different G protein-dependent effectors.

The mating factor receptors of yeast cells (STE2 and STE3) span the membrane of the yeast cell seven times and are coupled to yeast G proteins. The GPA1, STE4, and STE18 products are the yeast homologues of the  $\alpha$ ,  $\beta$  and  $\gamma$  subunits of mammalian G proteins, respectively. (Nakafuku et al. 1987. Proc. Natl. Acad. Sci USA 84:2140; Whiteway et al. 1989. Cell. 56:467). Since yeast cells have GPCRs analogous to those found in mammalian cells, experiments have also been undertaken to express functional GPCRs in yeast cells. The use of yeast cells for such expression provides advantages in terms of the ease of manipulating the cells, but also presents particular challenges in achieving efficient coupling and functional integration.

Methods for improving the functional integration of heterologous GPCRs in yeast cells are still needed. The prior art teaches numerous examples of the expression of heterologous receptors in yeast cells which fail to couple to yeast G proteins and thus are not functionally integrated into a yeast signaling pathway. For example Huang et al. teach that the rat M5 receptor, when expressed in yeast cells did not couple to the pheromone response pathway (Huang et al. Biochem. and Biophys. Res. Comm. 1992. 182:1180).

Previous work has demonstrated coupling of heterologous receptors in yeast by the expression of entire foreign G protein subunits in yeast cells (U.S. Patent 5,482,835 to King et al.). Another approach was taken by Kang (1990 Mol. Cell. Biol. 10:2582-2590) who made GPA1-G $\alpha$  chimeric subunits which comprised large portions, e.g., over 300 amino acids of GPA1. However, the chimeras made by Kang et al. were

- 3 -

assayed for their ability to complement a gpa1 null phenotype (i.e., constitutive activation of the pheromone response pathway) in *S. cerevisiae*. a situation in which it was desirable to retain a substantial portion of the GPA sequence. Clearly, a method for optimizing the functional integration of a heterologous GPCR into a signaling pathway  
5 in a yeast cell expressing such a receptor would be of great value in developing assays to identify receptor agonists and antagonists.

### Summary of the Invention

The present invention provides an important advance in drug screening  
10 methodologies previously known in the art by providing, *inter alia*, a means by which expression of heterologously expressed receptors is enhanced and a means by which coupling of heterologously expressed GPCRs to G protein subunits is enhanced.

The present invention pertains to novel yeast cells which are useful for the expression of functional heterologous GPCRs. In certain embodiments, the subject  
15 yeast cells comprise modified and/or heterologous G protein subunits which enhance the functional integration of heterologous GPCRs into a yeast signaling pathway. The modified G protein subunits can be altered by mutation and/or can be chimeric, i.e., can comprise a polypeptide derived from a yeast G protein subunit and one or more polypeptides derived from heterologous G protein subunits. The yeast cells of the  
20 present invention can be used in novel screening assays which can be used to screen for modulators of GPCRs.

In one aspect, the invention is a recombinant yeast cell which comprises:  
a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is  
25 modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast pheromone receptor in a pheromone response pathway of the yeast cell; and  
a chimeric G protein subunit selected from the group consisting of:  
a non-naturally occurring G protein subunit which comprises a  
30 sequence from a heterologous G protein subunit in which at least one amino acid substitution has been introduced compared to the wild type sequence;  
an endogenous STE 18 subunit operably linked to a polypeptide of a heterologous G<sub>y</sub> subunit;  
an endogenous yeast Gpa1 subunit in which at least the last four  
35 C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit; and

an endogenous yeast G<sub>p</sub>a1 subunit in which at least the last four C-terminal amino acids of said G<sub>p</sub>a1 are replaced with at least the last four C-terminal amino acids of a first heterologous G protein subunit, and in which the N-terminus of said G<sub>p</sub>a1 is operably linked to at least the first five N-terminal amino acids of a second 5 heterologous G protein subunit, wherein said first and second heterologous G protein subunits are the same or different: such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone system pathway of said yeast cell; and  
wherein modulation of the signal transduction activity of said heterologous GPCR by an 10 extracellular signal provides a detectable signal.

In one aspect, the invention provides a yeast cell comprising a heterologous G protein coupled receptor; and a non-naturally occurring G protein subunit which has a sequence from a heterologous G protein subunit, but in which at least one amino acid substitution has been introduced compared to the wild type 15 sequence. The expression of the non-naturally occurring G protein subunit functionally integrates the heterologous G protein coupled receptor into the yeast cell pheromone signaling pathway.

In one embodiment, the yeast cell of the present invention has a non-naturally occurring G protein subunit which is a mutant mammalian G<sub>a</sub> subunit. In 20 preferred embodiments, the mutant mammalian G<sub>a</sub> subunit comprises a sequence from a heterologous G protein subunit having a mutation selected from the group consisting of: G<sub>a</sub>16(S270P); G<sub>a</sub>s(D229S); G<sub>a</sub>s(D229V); G<sub>a</sub>s(N254D); G<sub>a</sub>s(S286P); G<sub>a</sub>s (E10K); G<sub>a</sub>i2-G<sub>a</sub>oB (S280P); G<sub>a</sub>12 (Q229L); G<sub>a</sub>12 (G228A); and G<sub>a</sub>i2 (S288P).

In another embodiment, the yeast cell of the present invention comprises 25 a non-naturally occurring G protein subunit which is a yeast-mammalian G protein subunit chimera comprising a first polypeptide from a yeast G protein subunit and a second polypeptide from a mutant mammalian G protein subunit. In preferred embodiments the second polypeptide of the chimera comprises a mutant mammalian G<sub>a</sub> subunit selected from the group consisting of: G<sub>a</sub>16(S270P); G<sub>a</sub>s(D229S); G<sub>a</sub> s(D229V); G<sub>a</sub>s(N254D); G<sub>a</sub>s(S286P); G<sub>a</sub>s (E10K); G<sub>a</sub>i2-G<sub>a</sub>oB (S280P); G<sub>a</sub>12 30 (Q229L); G<sub>a</sub>12 (G228A); and G<sub>a</sub>i2 (S288P).

In another preferred embodiment, a yeast cell of the present invention has 35 a chimeric G protein subunit, which comprises a first polypeptide from a yeast G protein subunit and a second polypeptide from a heterologous G protein subunit, where the first polypeptide is selected from the group consisting of: a polypeptide comprising about 40

- 5 -

amino acids from the amino terminus of yeast GPA1; and a polypeptide from yeast STE 18.

In yet another preferred embodiment, a yeast cell of the present invention has a chimeric G protein subunit where the first polypeptide of the chimera comprises 5 about 40 amino acids from the amino terminus of yeast GPA1 and said second polypeptide of the chimera is from a heterologous G protein  $\alpha$  subunit.

In yet another embodiment a yeast cell of the present invention has a chimeric G protein subunit in which the first polypeptide is from yeast STE 18 and the second polypeptide is from a heterologous G protein  $\gamma$  subunit.

10 In preferred embodiments, a heterologous G protein subunit of the present invention is mammalian. In particularly preferred embodiments, a heterologous G protein subunit of the present invention is human.

15 In one embodiment, a yeast cell of the present invention comprises a chimeric G protein subunit in which at least one of the first and second polypeptides comprises a naturally occurring amino acid sequence. In still another embodiment, at least one of the first and second polypeptides of the chimeric G protein subunit comprises a non-naturally occurring amino acid sequence.

20 In preferred embodiments, a yeast cell of the present invention comprises a heterologous G protein coupled receptor which is functionally integrated into the yeast cell.

In particularly preferred embodiments, a modified or chimeric G protein subunit of the present invention demonstrates enhanced coupling to a heterologous G protein coupled receptor expressed by a yeast cell when compared to that demonstrated by an endogenous yeast G protein subunit.

25 In one embodiment a yeast cell of the present invention comprises a chimeric G protein subunit in which the second polypeptide is from the human  $G\gamma 2$  subunit. In a preferred embodiment, the second polypeptide comprises the amino acid sequence Arg Glu Lys Lys Phe Phe (amino acids 19-24 of SEQ ID NO: 33). In a particularly preferred embodiment, the chimeric G protein subunit comprises the 30 sequence shown in SEQ ID NO: 33.

In a preferred embodiment, a yeast cell of the present invention comprises a chimeric G protein subunit selected from the group consisting of: gp $\alpha$ 1 (41)-G $\alpha$ i2; gp $\alpha$ 1 (41)-G $\alpha$ 16; and gp $\alpha$ 1 (41)-G $\alpha$ s. In a more preferred embodiment, a yeast cell of the present invention comprises a chimeric G protein subunit in which the G $\alpha$ i2, 35 G $\alpha$ 16, or G $\alpha$ s portion of the chimeric G protein subunit comprises an amino acid substitution compared to wild type G $\alpha$ i2, G $\alpha$ 16, or G $\alpha$ s.

In yet another embodiment, a yeast cell of the present invention comprises a second chimeric G protein subunit, in which the second chimeric G protein subunit has a first polypeptide from a yeast G protein subunit and a second polypeptide from a mammalian G protein subunit, and wherein the second chimeric G protein 5 subunit is different from a first chimeric G protein subunit expressed by the yeast cell. In a preferred embodiment, the second polypeptide of the second chimeric G protein subunit is from a protein selected from the group consisting of: a mammalian G $\alpha$  subunit, a mammalian G $\beta$  subunit, and a mammalian G $\gamma$  subunit.

10 In preferred embodiments, a yeast cell of the present invention does not produce an endogenous yeast pheromone system receptor protein in functional form.

In certain embodiments, a yeast cell of the present invention comprises a indicator gene that produces a detectable signal upon functional coupling of the heterologous G protein coupled receptor to the G protein.

15 In preferred embodiments, a yeast cell of the present invention comprises a heterologous G protein coupled receptor which is an orphan receptor.

In another embodiment the invention provides an assay to identify compounds capable of modulating the dissociation of G $\alpha$  and G $\beta\gamma$ , comprising the steps of: providing a yeast cell which comprises a heterologous G protein coupled receptor, a modified G protein subunit, and an indicator gene, contacting the yeast with a test 20 compound; and identifying compounds which induce a change in a detectable signal in the yeast cell, wherein said detectable signal indicates dissociation of G $\alpha$  and G $\beta\gamma$ .

In certain embodiments, an assay of the present invention is used to test compounds from a library of non-peptidic organic molecules.

25 In another aspect, the invention provides a method for identifying a compound which modulates a heterologous G protein coupled receptor, comprising: providing a first, second, third, and fourth yeast cell, each cell comprising a G protein, wherein:

- 1) the first yeast cell comprises a first chimeric G protein subunit comprising a first polypeptide from a yeast G protein subunit and a second polypeptide 30 from a mammalian G protein subunit; 2) the second yeast cell comprises a second chimeric G protein subunit comprising a first polypeptide derived from a yeast G protein subunit and a second polypeptide from a mammalian G protein subunit, the second chimeric G protein subunit being different from said first chimeric G protein subunit; 3) the third yeast cell comprises a third chimeric G protein subunit comprising a first 35 polypeptide from a yeast G protein subunit and a second polypeptide from a mammalian G protein subunit, the third chimeric G protein subunit being different from said first

and second chimeric G protein subunits; 4) the fourth yeast cell comprises an endogenous yeast gpa1 G protein subunit; and an expressible gene construct encoding a heterologous G protein coupled receptor (GPCR) which couples to the yeast pheromone system pathway; and an indicator gene that produces a detectable signal upon functional coupling of the heterologous G protein coupled receptor to the G protein. The assay 5 comprises contacting the first, second, third, and fourth yeast cells with a test compound; and determining whether the test compound induces a change in a detectable signal in at least one of the first, second, third, or fourth yeast cells to thereby identify a compound which modulates a heterologous GPCR.

10 In preferred embodiments, the assay is performed using a yeast cell which comprises a chimeric G protein subunit in which the second polypeptide is from a mammalian G $\alpha$ i subunit. In another preferred embodiment, the second polypeptide of the chimeric G protein subunit is from a mammalian G $\alpha$ 16 subunit. In yet another preferred embodiment, the second polypeptide of the chimeric G protein subunit is from 15 a mammalian G $\alpha$ s subunit.

In yet another embodiment of the assay, the first chimeric G protein subunit comprises a polypeptide from mammalian G $\alpha$ 12, the second chimeric G protein subunit comprises a polypeptide from mammalian G $\alpha$ 16, and the third chimeric G protein subunit comprises a polypeptide from mammalian G $\alpha$ s. In a preferred 20 embodiment, the second chimeric G protein subunit comprises G $\alpha$ 16(S270P) and the third chimeric G protein subunit comprises G $\alpha$ s(D229S).

25 In another preferred embodiment, each of the first, second, and third yeast cells further comprises a fourth chimeric G protein subunit, said fourth chimeric G protein subunit comprising a first polypeptide from yeast STE 18 and a second polypeptide from a mammalian G protein  $\gamma$  subunit.

In certain embodiments of the subject assays, the first, second, third, and fourth yeast cells are contacted with each member of a library of test compounds. In a preferred embodiment, each member of said library is a non-peptidic organic molecule.

30 In a preferred embodiment, the first, second, third, and fourth yeast cells are *Saccharomyces cerevisiae* cells.

35 In one embodiment of the invention, the indicator gene that gives rise to a detectable signal is selected from the group consisting of:  $\beta$  galactosidase, alkaline phosphatase, horseradish peroxidase, exoglucanase, luciferase, BAR1, PHO5, green fluorescent protein and chloramphenicol acetyl transferase.

In yet another embodiment, the indicator gene that gives rise to a detectable signal is a HIS 3 gene.

In one embodiment of the invention, the heterologous G protein coupled receptor which is expressed by a yeast cell is an orphan receptor.

In still yet another aspect, the invention is directed to carboxy terminal chimeric G protein subunits, and sandwich chimeric G protein subunits. Thus, in one embodiment, a yeast cell of the invention includes a chimeric G protein subunit comprising an endogenous yeast G<sub>p</sub>a1 subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit.

In another embodiment, a yeast cell of the invention includes a chimeric G protein subunit comprises an endogenous yeast G<sub>p</sub>a1 subunit in which at least the last four C-terminal amino acids of said G<sub>p</sub>a1 are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, and in which the N-terminus of said G<sub>p</sub>a1 is operably linked to at least the first five N-terminal amino acids of a heterologous G protein subunit, wherein said heterologous G protein subunits are the same or different.

In another embodiment, a yeast cell of the invention which comprises:  
a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast pheromone receptor in a pheromone response pathway of the yeast cell; and  
a chimeric G protein subunit comprising an endogenous yeast G<sub>p</sub>a1 subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell; and wherein modulation of the signal transduction activity of said heterologous GPCR by an extracellular signal provides a detectable signal.

In another embodiment, the yeast cell comprises a heterologous polypeptide, wherein the heterologous polypeptide is transported to a location allowing interaction with the extracellular region of the receptor expressed in the cell membrane; and wherein the heterologous polypeptide is expressed at a sufficient level such that modulation of the signal transduction activity of the receptor by the heterologous polypeptide provides a detectable signal.

In yet another embodiment, the yeast cell comprises a reporter construct that is activated by the pheromone response pathway, wherein the heterologous

- 9 -

polypeptide is expressed at a sufficient level such that modulation of the signal transduction activity of the receptor by the heterologous polypeptide provides a detectable signal mediated by the reporter construct.

In still another embodiment, a yeast cell of the invention comprises a mutation in at least one gene selected from the group consisting of *FAR1*, *SST2*, *BARI*, *SVG1*, *STE2*, *STE3*, *STE14*, *MFa1*, *MFa2*, *MFa1* and *MFa2*; and wherein modulation of the signal transduction activity of said heterologous GPCR by an extracellular signal provides a detectable signal.

The invention is also directed to a mixture of recombinant yeast cells, each cell of which has a cell membrane and each cell of which comprises:

a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal;

a chimeric G protein subunit comprising an endogenous yeast G<sub>pa1</sub> subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell; and

a heterologous polypeptide, wherein the heterologous polypeptide is transported to a location allowing interaction with the extracellular region of said receptor expressed in the cell membrane; wherein collectively the mixture of cells expresses a library of said heterologous polypeptides, said library being expressible at a sufficient level such that modulation of the signal transduction activity of said receptor by a heterologous polypeptide within the library provides a detectable signal.

In another aspect the invention is directed to sandwich chimeric G protein subunits. In one embodiment, a yeast cell of the invention comprises:

a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast pheromone receptor in a pheromone response pathway of the yeast cell; and

a chimeric G protein subunit comprising an endogenous yeast G<sub>pa1</sub> subunit in which at least the last four C-terminal amino acids of said G<sub>pa1</sub> are replaced with at least the last four C-terminal amino acids of a first heterologous G protein subunit, and in which the N-terminus of said G<sub>pa1</sub> is operably linked to at least

- 10 -

the first five N-terminal amino acids of a second heterologous G protein subunit.  
wherein said first and second heterologous G protein subunits are the same or different,  
such that expression of said chimeric G protein subunit functionally integrates said  
heterologous GPCR into the pheromone response pathway of said yeast cell; and  
5 wherein modulation of the signal transduction activity of said heterologous GPCR by an  
extracellular signal provides a detectable signal.

In another embodiment, the invention is a mixture of recombinant yeast  
cells, each cell of which has a cell membrane and each cell of which comprises:

10 a heterologous G protein-coupled receptor (GPCR) expressed in  
the cell membrane of said yeast cell such that signal transduction activity via said  
receptor is modulated by interaction of an extracellular region of the receptor with an  
extracellular signal;

15 a chimeric G protein subunit comprising an endogenous yeast  
Gpa1 subunit in which at least the last four C-terminal amino acids of said Gpa1 are  
replaced with at least the last four C-terminal amino acids of a heterologous G protein  
subunit, and in which the N-terminus of said Gpa1 is operably linked to at least the first  
five N-terminal amino acids of a heterologous G protein subunit, such that expression of  
said chimeric G protein subunit functionally integrates said heterologous GPCR into the  
pheromone response pathway of said yeast cell; and

20 a heterologous polypeptide, wherein the heterologous polypeptide  
is transported to a location allowing interaction with the extracellular region of said  
receptor expressed in the cell membrane; wherein collectively the mixture of cells  
expresses a library of said heterologous polypeptides, said library being expressible at a  
sufficient level such that modulation of the signal transduction activity of said receptor  
25 by a heterologous polypeptide within the library provides a detectable signal.

In another embodiment, a yeast cell of the invention comprises:

30 a heterologous G protein-coupled receptor (GPCR) expressed in  
the cell membrane of said yeast cell such that signal transduction activity via said  
receptor is modulated by interaction of an extracellular region of the receptor with an  
extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous  
yeast pheromone receptor in a pheromone response pathway of the yeast cell; and

35 a chimeric G protein subunit comprising an endogenous yeast  
Gpa1 subunit in which at least the last four C-terminal amino acids of said Gpa1 are  
replaced with at least the last four C-terminal amino acids of a first heterologous G  
protein subunit, and in which at least the first four N-terminal amino acids of said Gpa1  
are replaced with at least the first six N-terminal amino acids of a second heterologous G

protein subunit, wherein said first and second heterologous G protein subunits are the same or different; such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell; and

- 5 wherein modulation of the signal transduction activity of said heterologous GPCR by an extracellular signal provides a detectable signal.

In a preferred embodiment, the chimeric G protein subunit comprises an endogenous yeast G<sub>p</sub>a1 subunit in which the last five C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last five C-terminal amino acids of a first heterologous G protein subunit, and in which the first five N-terminal amino acids of said G<sub>p</sub>a1 are replaced with the first 11 N-terminal amino acids of a second heterologous G protein subunit, wherein said first and second heterologous G protein subunits are the same.

In another preferred embodiment, the chimeric G protein subunit comprises an endogenous yeast G<sub>p</sub>a1 subunit in which the last five C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last five C-terminal amino acids of a first heterologous G protein subunit, and in which the first 21 N-terminal amino acids of said G<sub>p</sub>a1 are replaced with the first 21 N-terminal amino acids of a second heterologous G protein subunit, wherein said first and second heterologous G protein subunits are the same.

20 In another aspect, the invention is a chimeric G-protein subunit which comprises an endogenous G<sub>p</sub>a1 subunit in which at least the last four C-terminal amino acids of said G<sub>p</sub>a1 are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit.

25 In a preferred embodiment, the last five C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last five C-terminal amino acids of a heterologous G protein subunit.

In another preferred embodiment, the last six C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last six C-terminal amino acids of a heterologous G protein subunit.

30 In another embodiment, the invention is a chimeric G-protein subunit which comprises an endogenous G<sub>p</sub>a1 subunit in which at least the last four C-terminal amino acids of said G<sub>p</sub>a1 are replaced with at least the last four C-terminal amino acids of a first heterologous G protein subunit, and in which the N-terminus of said G<sub>p</sub>a1 is operably linked to at least the first five N-terminal amino acids of a second heterologous G protein subunit, wherein said first and second heterologous G protein subunits are the same or different.

In a preferred embodiment, the last five C-terminal amino acids of said G<sub>pα1</sub> are replaced with the last five C-terminal amino acids of said first heterologous G-protein subunit, and in which the first five N-terminal amino acids of said G<sub>pα1</sub> are replaced with the first 11 N-terminal amino acids of said second heterologous G protein subunit.

5 In another preferred embodiment, the last five C-terminal amino acids of said G<sub>pα1</sub> are replaced with the last five C-terminal amino acids of said first heterologous G-protein subunit, and in which the first 21 N-terminal amino acids of said G<sub>pα1</sub> are replaced with the first 21 N-terminal amino acids of said second heterologous G protein subunit.

10 In another preferred embodiment, the first and second heterologous G protein subunits are the same.

In another aspect, the invention is a method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

15 contacting a mixture of yeast cells of the invention with a test compound;

allowing cells within the mixture to generate a detectable signal; and

identifying the test compound as a modulator of said receptor.

20 In another embodiment, the invention is method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

contacting a mixture of yeast cells of the invention with a ligand of said receptor;

25 allowing cells within the mixture to generate a detectable signal; and

identifying a heterologous polypeptide expressed by the yeast cells as a modulator of said receptor.

30 In another embodiment, the invention is a method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

contacting a first mixture of yeast cells as claimed in claim 78 with a second mixture of yeast cells, wherein collectively the second mixture of yeast cells expresses a library of heterologous test polypeptides that are transported to a

35 location allowing interaction with the extracellular region of said receptor expressed in the cell membrane of the yeast cells of the first mixture;

- 13 -

allowing cells within the first mixture to generate a detectable signal; and

identifying a heterologous test polypeptide within the second mixture as a modulator of said receptor expressed by the yeast cell.

5 In another embodiment, a method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

providing a mixture of recombinant yeast cells of the invention wherein a heterologous polypeptide in each of said yeast cells is a heterologous test polypeptide;

10 allowing cells within the mixture to generate a detectable signal; and

identifying the heterologous test polypeptide as a modulator of said heterologous receptor.

15 In still another embodiment, a method is provided for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell of the invention, comprising:

contacting a mixture of yeast cells of the invention with a ligand of said receptor;

20 allowing cells within the mixture to generate a detectable signal; and

identifying a heterologous polypeptide expressed by the yeast cells as a modulator of said receptor.

The invention is also directed to the following chimeras: the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 107; the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 108; the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 109; the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 110; the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 111; the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 112; the chimeric G 25 protein subunit comprising the amino acid sequence of SEQ ID NO: 113; the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 114; the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 115; the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 118; and the chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 123.

### Brief Description of the Drawings

Figure 1 is a sequence alignment of N-terminal regions of G $\alpha$  subunits and N-terminal sequences of GPA41-G $\alpha$  hybrid proteins.

5 Figure 2 is an illustration of the effect of mutations to a mammalian G protein subunit, G $\alpha$ s, on the coupling of the mammalian G $\alpha$ s subunit to yeast G $\beta$  $\gamma$ .

### Detailed Description of the Invention

The present invention provides, inter alia, rapid, effective assays for screening and identifying pharmaceutically effective compounds that specifically modulate the activity of a heterologous G protein coupled receptor (GPCR) expressed in a yeast cell. The subject assays enable rapid screening of large numbers of compounds (e.g., compounds in a library) to identify those which are receptor agonists or antagonists. Compositions of matter, such as novel recombinant yeast cells and novel gene constructs are also embraced by the present invention. The instant assays provide a convenient format for discovering compounds which can be useful to modulate cellular function, as well as to understand the pharmacology of compounds that specifically interact with cellular receptors.

In the practice of the instant invention, standard techniques known in the art can be used. See for example, Sherman. 1991. Methods Enzymol. 194:3; Sherman and Hicks. 1991. Methods Enzymol. 194:21; Sambrook et al. Molecular Cloning A Laboratory Manual, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989 or 1991 edition.

Before further description of the invention, certain terms employed in the specification, examples and appended claims are, for convenience, collected here.

#### I. Definitions

The term "compound" as used herein (e.g., as in "test compound") is meant to include both exogenously added test compounds and peptides endogenously expressed from a peptide library. For example, in certain embodiments, the reagent cell also produces the test compound which is being screened. For instance, the reagent cell can produce, e.g., a test polypeptide, a test nucleic acid and/or a test carbohydrate which is screened for its ability to modulate the heterologous receptor activity. In such embodiments, a culture of such reagent cells will collectively provide a library of potential effector molecules and those members of the library which either agonize or antagonize the receptor function can be selected and identified. Moreover, it will be

apparent that the reagent cell can be used to detect agents which transduce a signal via the receptor of interest.

In other embodiments, the test compound is exogenously added. In such embodiments the test compound is contacted with the reagent cell. Exemplary 5 compounds which can be screened for activity include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, and natural product extract libraries. In such embodiments, both compounds which agonize or antagonize the receptor mediated signaling function can be selected and identified.

The term "non-peptidic compound" is intended to encompass compounds 10 that are comprised, at least in part, of molecular structures different from naturally-occurring L-amino acid residues linked by natural peptide bonds. However, "non-peptidic compounds" are intended to include compounds composed, in whole or in part, of peptidomimetic structures, such as D-amino acids, non-naturally-occurring L-amino acids, modified peptide backbones and the like, as well as compounds that are 15 composed, in whole or in part, of molecular structures unrelated to naturally-occurring L-amino acid residues linked by natural peptide bonds. "Non-peptidic compounds" also are intended to include natural products.

As used herein, "recombinant cells" include any cells that have been modified by the introduction of heterologous DNA. Control cells include cells that are 20 substantially identical to the recombinant cells, but do not express one or more of the proteins encoded by the heterologous DNA, e.g., do not include or express a reporter gene construct, receptor or test polypeptide.

As used herein, "heterologous DNA" or "heterologous nucleic acid" includes DNA that does not occur naturally as part of the genome in which it is present 25 or which is found in a location or locations in the genome that differs from that in which it occurs in nature. Heterologous DNA is not naturally occurring in that position or is not endogenous to the cell into which it is introduced, but has been obtained from another cell. Generally, although not necessarily, such DNA encodes proteins that are not normally produced by the cell in which it is expressed. Heterologous DNA can be 30 from the same species, although in preferred embodiments, it is from a different species. In particularly preferred embodiments, it is mammalian, e.g., human. Heterologous DNA may also be referred to as foreign DNA. Any DNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which it is expressed is herein encompassed by the term heterologous DNA. Examples of heterologous DNA 35 include, but are not limited to, DNA that encodes test polypeptides, receptors, reporter

genes, transcriptional and translational regulatory sequences, or selectable or traceable marker proteins, such as a protein that confers drug resistance.

The terms "heterologous protein", "recombinant protein", and "exogenous protein" are used interchangeably throughout the specification and refer to a polypeptide which is produced by recombinant DNA techniques, wherein generally, 5 DNA encoding the polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. That is, the polypeptide is expressed from a heterologous nucleic acid.

"Orphan receptors" is a designation given to a receptors for which no 10 specific natural ligand has yet been described.

As used herein, the term "extracellular signal" is intended to encompass molecules and changes in the environment that are transduced intracellularly via cell surface proteins that interact, directly or indirectly, with the extracellular signal. An extracellular signal or effector molecule includes any compound or substance that in 15 some manner alters the activity of a cell surface protein. Examples of such signals include, but are not limited to, molecules such as acetylcholine, growth factors and hormones, lipids, sugars and nucleotides that bind to cell surface receptors and modulate the activity of such receptors. The term, "extracellular signal" also includes as yet unidentified substances that modulate the activity of a cellular receptor, and thereby 20 influence intracellular functions. Such extracellular signals are potential pharmacological agents that may be used to treat specific diseases by modulating the activity of specific cell surface receptors.

As used herein, "cell surface receptor" refers to molecules that occur on the surface of cells, interact with the extracellular environment, and transmit or 25 transduce the information regarding the environment intracellularly in a manner that may modulate intracellular second messenger activities or transcription of specific promoters, resulting in transcription of specific genes. A "heterologous receptor" is a specific embodiment of a "heterologous protein", wherein the heterologous receptor is encoded by heterologous DNA and, upon expression of this heterologous DNA in a 30 recombinant cell, the heterologous receptor is expressed in the recombinant cell. Preferred receptors are G protein coupled receptors and exemplary GPCRs are described in detail herein.

The term "signal transduction" is intended to encompass the processing of physical or chemical signals from the extracellular environment through the cell 35 membrane and into the cell, and may occur through one or more of several mechanisms, such as activation/inactivation of enzymes (such as proteases, or other enzymes which

may alter phosphorylation patterns or other post-translational modifications), activation of ion channels or intracellular ion stores, effector enzyme activation via guanine nucleotide binding protein intermediates, formation of inositol phosphate, activation or inactivation of adenylyl cyclase, direct activation (or inhibition) of a transcriptional factor and/or activation. A "signaling pathway" refers to the components involved in "signal transduction" of a particular signal into a cell. The term "endogenous signaling pathway" indicates that some or all of the components of the signaling pathway are naturally-occurring components of the cell. An example of such a pathway is the endogenous pheromone system pathway of yeast.

The term "functionally integrated" (as in a receptor that is "functionally integrated into a signaling pathway in a cell" or "functionally integrated into an endogenous yeast signaling pathway") is intended to refer to the ability of the receptor to be expressed at the surface of the cell and the ability of the expressed receptor to bind to modulators (e.g., a ligand of the receptor) and transduce signals into the cell via components of a signaling pathway of the cell. For example, a G protein coupled receptor (GPCR) which is functionally integrated into an endogenous pheromone response pathway of a yeast cell is expressed on the surface of the yeast cell, couples to a G protein of the pheromone response pathway within the yeast cell and transduces a signal in that yeast cell upon binding of a modulator to the receptor. For a G protein subunit to be functionally integrated into a yeast cell such a subunit, e.g., a chimeric, mutant or heterologous subunit, must be capable of coupling both to the GPCR and to the other G protein subunits, which can also be endogenous to the yeast cell, can be chimeric, or can be heterologous. A transduced signal may be detected by measuring any one of a number of responses to mating factors which occur in a yeast cell, e.g., growth arrest or transcription of an indicator gene responsive to signals produced by modulation of a pheromone system pathway.

The term "indicator gene" generically refers to an expressible (e.g., able to transcribed and (optionally) translated) DNA sequence which is expressed in response to a signal transduction pathway modulated by a target receptor or ion channel.

Exemplary indicator genes include unmodified endogenous genes of the host cell, modified endogenous genes, or a reporter gene of a heterologous construct, e.g., as part of a reporter gene construct.

The term "endogenous gene" is intended to refer to a gene in a cell that is naturally part of the genome of the cell and which, most preferably, is present in its natural location in the genome (as opposed to "heterologous" DNA which has been

introduced into the cell). Likewise, the term "endogenous protein" is intended to include proteins of a cell that are encoded by endogenous genes of the cell.

An endogenous gene that is to be used as an indicator gene may comprise the natural regulatory elements of the gene (e.g., the native promoter/enhancer elements

5 that naturally regulate expression of the gene) or the endogenous gene can be "operatively linked to" (i.e., functionally coupled to) a "heterologous promoter" (or other heterologous regulatory elements). A "heterologous promoter" refers to a promoter that does not naturally regulate the gene to which the heterologous promoter is operatively linked. For example, an endogenous yeast gene that is not normally pheromone-

10 responsive can be operatively linked to a heterologous promoter that is responsive to signals produced by the yeast pheromone system to thereby confer pheromone responsiveness on the endogenous yeast gene. Methods of using endogenous yeast genes as indicator genes are described further in USSN \_\_\_\_\_, entitled, "Methods and Compositions for Identifying Receptor Effectors", filed on September 24, 1997

15 (Attorney Docket No. CPI-031CP2) the contents of which are hereby expressly incorporated herein by this reference.

The term "detecting an alteration in a signal produced by an endogenous signaling pathway" (e.g., an endogenous yeast signaling pathway) is intended to encompass the detection of alterations in endogenous second messengers produced upon 20 activation of components of the endogenous signaling pathway, alterations in endogenous gene transcription induced upon activation of components of the endogenous signaling pathway, and/or alterations in the activity of an endogenous protein(s) upon activation of components of the endogenous signaling pathway. In certain embodiments, the term "detecting an alteration in a signal produced by an 25 endogenous signaling pathway" can also encompass assaying general, global changes to the cell such as changes in cell growth or cell morphology.

As used herein, a "reporter gene construct" refers to a nucleic acid that includes a "reporter gene" operatively linked to a transcriptional regulatory sequences. Transcription of the reporter gene is controlled by these sequences. The activity of at 30 least one or more of these control sequences is directly or indirectly regulated by the target receptor protein. The transcriptional regulatory sequences include the promoter and other regulatory regions, such as enhancer sequences, that modulate the activity of the promoter, or regulatory sequences that modulate the activity or efficiency of the RNA polymerase that recognizes the promoter, or regulatory sequences which are 35 recognized by effector molecules, including those that are specifically induced by interaction of an extracellular signal with the target receptor. For example, modulation

of the activity of the promoter may be effected by altering the RNA polymerase binding to the promoter region, or, alternatively, by interfering with initiation of transcription or elongation of the mRNA. Such sequences are herein collectively referred to as transcriptional regulatory elements or sequences. In addition, the construct may include 5 sequences of nucleotides that alter translation of the resulting mRNA, thereby altering the amount of reporter gene product. The reporter gene constructs of the present invention provide a detectable readout in response to signals transduced in response to modulation of a heterologously expressed receptor.

The term "modulation", as in "modulation of a (heterologous) receptor" 10 and "modulation of a signal transduction activity of a receptor protein" is intended to encompass, in its various grammatical forms, induction and/or potentiation, as well as inhibition and/or downregulation of receptor activity and/or one or more signal transduction pathways downstream of a receptor.

Agonists and antagonists are "receptor effector" molecules that modulate 15 signal transduction via a receptor. Receptor effector molecules are capable of binding to the receptor, though not necessarily at the binding site of the natural ligand or otherwise modulating the activity of the receptor, for example, by influencing the activity of components which regulate the receptor, or which function in the signal transduction pathway initiated by the receptor. Receptor effectors can modulate signal transduction 20 when used alone, i.e. can be surrogate ligands, or can alter signal transduction in the presence of the natural ligand or other known activators, either to enhance or inhibit signaling by the natural ligand. For example, "antagonists" are molecules that block or decrease the signal transduction activity of receptor, e.g., they can competitively, noncompetitively, and/or allosterically inhibit signal transduction from the receptor, 25 whereas "agonists" potentiate, induce or otherwise enhance the signal transduction activity of a receptor. The term "surrogate ligand" refers to an agonist which induces signal transduction from a receptor.

The term "autocrine cell", as used herein, refers to a cell which produces 30 a substance which can stimulate a receptor located on or within the same cell as that which produces the substance. For example, wild-type yeast MAT $\alpha$  and MAT $\alpha$  cells are not autocrine. However, a yeast cell which produces both  $\alpha$ -factor and  $\alpha$ -factor receptor, or both  $a$ -factor and  $a$ -factor receptor, in functional form, is autocrine. By extension, cells which produce a peptide which is being screened for the ability to activate a receptor (e.g., by activating a G protein-coupled receptor) and also express the receptor 35 are called "autocrine cells". In some instances, such cells can also be referred to as "putative autocrine cells" since some of the cells will express peptides from the library

which will not activate the receptor which is expressed. In a library of such cells, in which a multitude of different peptides are produced, it is likely that one or more of the cells will be "autocrine" in the stricter sense of the term.

As used herein the term "chimeric" G protein subunit refers to a G protein subunit composed of at least two discrete polypeptides, a first polypeptide from a yeast G protein subunit and a second polypeptide from a heterologous G protein subunit. Each of the first and second polypeptides are encoded by a nucleic acid construct and are operatively linked such that upon expression of the construct, a functional chimeric G protein subunit is produced, i.e., a fusion protein comprising the first polypeptide linked to the second polypeptide. In preferred embodiment, the heterologous G protein subunit is mammalian. In particularly preferred embodiments, the heterologous G protein subunit is human. For example, chimeric G protein subunits of the present invention can comprise a polypeptide from GPA1 linked to G $\alpha$ , STE18, linked to G $\gamma$ , or STE4, linked to G $\beta$ . In preferred embodiments, in particular for chimeric G $\alpha$  subunits, the portion of the chimeric subunit from yeast GPA1 comprises a portion of the amino terminus of GPA1 and is less than 330 amino acids in length. In particularly preferred embodiments, the portion of the chimeric subunit derived from GPA1 is about 40 amino acids. In another embodiment, the portion of the chimeric subunit derived from GPA1 is about 20 amino acids.

As used herein, the term "not produced in functional form" with regard to endogenous yeast proteins is intended to encompass proteins which are not produced in functional form for any number of reasons, for example, because of a mutation to the gene which encodes the protein or a deletion, e.g., a disruption, of the gene which encodes the protein. The term "not produced in functional form" is also intended to include conditional mutations (e.g. temperature sensitive mutation, wherein the protein is not produced in functional form under certain conditions).

As used herein, the terms used to indicate amino acid mutations, such as "S270P" and the like, represent the wild type amino acid residue (in standard one letter code), followed by the amino acid position, followed by the substituted amino acid (in standard one letter code). Thus, S270P indicates substitution of the wild type serine at position 270 with proline. The terms such as "G $\alpha$ s(S270P)" and the like represent the G protein having the indicated substitution. Thus, the term G $\alpha$ s(S270P) represents a G $\alpha$ s subunit having a proline substituted for the wild type serine at position 270.

The phrase "last e.g. four C-terminal amino acids" and "first, e.g., 6 N-terminal amino acids" as used herein with reference to polypeptides wherein the amino acids of the polypeptides are read in sequential order from left to right, with the N-

- 21 -

terminus of the polypeptide being at the far left and the C-terminus being at the far right. For example, in a polypeptide of 50 amino acids, the "last five C-terminal amino acids" would refer to amino acids 46 through 50. Similarly, the "first six N-terminal amino acids" would refer to amino acids 1 through 6.

5 The term "transport to a location", as used herein refers to a heterologous polypeptide which is produced in one region of the yeast cell but which moves to another region of the yeast cell.

10 The term "expressed at a sufficient level", as used herein refers to a heterologous polypeptide which is produced in an amount capable of modulating signal transduction activity of a receptor.

15 The term "first heterologous G protein subunit" is used interchangeably with the term "first heterologous G  $\alpha$  subunit", and refers to a member of the family of G  $\alpha$  subunits that is used to replace the C-terminal or N-terminal amino acids of the G $\alpha$ 1 protein. The "first heterologous G protein subunit" may also be operably linked to the N-terminal of the G $\alpha$ 1 protein. Examples of G  $\alpha$  subunits include, G $\alpha$ 16; G $\alpha$ s; G $\alpha$ s; G $\alpha$ s; G $\alpha$ s; G $\alpha$ i2; G $\alpha$ B; G $\alpha$ 12; G $\alpha$ 12; and G $\alpha$ i2. The term "second heterologous G protein subunit" is used interchangeably with the term "second heterologous G  $\alpha$  subunit", and refers to a member of the family of G  $\alpha$  subunits that is used to replace the C-terminal or N-terminal amino acids of the G $\alpha$ 1 protein. In 20 chimeric G proteins, the G $\alpha$ 1 protein can be replaced with a first heterologous G protein subunit at the C-terminal of the G $\alpha$ 1 protein. In "sandwich chimera G proteins", the C-terminal amino acid of the G $\alpha$ 1 protein can be replaced with a first heterologous G protein subunit, and the N-terminal can be replaced with, or operably linked to, a second heterologous G protein subunit. The first and second heterologous G 25 protein subunits can be the same, for example, Gq-G $\alpha$ 1-Gq, Gs-G $\alpha$ 1-Gs and the like, or the first and second heterologous G protein subunits can be different, for example Gq-G $\alpha$ 1-Gs, G $\alpha$ -G $\alpha$ 1-Gq, and the like.

With regard to polypeptides, the terms "operatively linked" and "operably linked", are used herein interchangeably and are intended to mean that two polypeptides 30 are connected in manner such that each polypeptide can serve its intended function. Typically, the two polypeptides are covalently attached through peptide bonds. The fusion protein is preferably produced by standard recombinant DNA techniques. For example, a DNA molecule encoding the first polypeptide is ligated to another DNA molecule encoding the second polypeptide, and the resultant hybrid DNA molecule is 35 expressed in a host cell to produce the fusion protein. The DNA molecules are ligated to each other in a 5' to 3' orientation such that, after ligation, the translational frame of the

- 22 -

encoded polypeptides is not altered (i.e., the DNA molecules are ligated to each other in-frame).

## II. General Overview of Assay

As set out above, the present invention relates to methods for identifying effectors of a receptor protein or receptor protein complex. The instant assays are characterized by the use of a mixture of recombinant yeast cells to sample test compounds for receptor agonists or antagonists. As described in greater detail below, the reagent cells express a heterologous GPCR protein functionally integrated into the cell and capable of transducing a detectable signal in the yeast cell. Exemplary GPCRs are discussed below. Compounds which either agonize or antagonize the receptor function can be selected and then identified based on biochemical signals produced by the receptor, or any more distal result of receptor-mediated stimulation, for example increases in endogenous mRNA expression, etc., or, in some embodiments, by the use of reporter genes responsive to such signals. In certain embodiments, the library of compounds to be tested is a library of peptides which is expressed by the yeast cells and causes stimulation in an autocrine fashion.

The ability of compounds to modulate the signal transduction activity of the target receptor can be scored for by detecting up or down-regulation of the detection signal. For example, GTPase activity, phospholipid hydrolysis, or protein phosphorylation stimulated by the receptor can be measured directly. Alternatively, the use of a reporter gene can provide a readout. In any event, a statistically significant change in the detection signal can be used to facilitate isolation of compounds of interest of those cells from the mixture which contain a nucleic acid encoding a test polypeptide which is an effector of the target receptor.

In certain embodiments, the yeast cells for use in the instant assays express heterologous GPCR and an endogenous G protein subunit which couples to that receptor. Preferably, the yeast cells of the present invention have been modified such that coupling of the GPCR to the yeast pheromone signaling pathway is enhanced. For example, in preferred embodiments, the yeast cells express a heterologous GPCR and mutated endogenous G protein subunit which facilitates functional integration of that receptor into the yeast cell. In another preferred embodiment, the yeast cells express a heterologous GPCR and a heterologous G protein subunit. In particularly preferred embodiments, the heterologous GPCR and the heterologous G protein subunit are of the same origin, e.g., mammalian. In yet another preferred embodiment, the yeast cells express a mutated heterologous G protein subunit.

In still another preferred embodiment, the yeast cells express a chimeric G protein subunit. In particularly preferred embodiments the heterologous GPCR and the heterologous segment of the chimeric G protein subunit are derived from the same source. In more preferred embodiments, the second amino acid sequence in the G 5 protein subunit chimera is derived from a mammalian G protein subunit. In particularly preferred embodiments, the second amino acid sequence is derived from a human G protein subunit sequence.

It will further be understood that the above embodiments are not mutually exclusive. For example, in certain preferred embodiments, a yeast cell may express a 10 first mutated or chimeric G protein subunit and a second, different mutated or chimeric G protein subunit to enhance coupling to the heterologous receptor.

In certain embodiments the yeast cells also express an indicator gene that produces a detectable signal upon functional coupling of the heterologous G protein coupled receptor to the G protein. In certain embodiments the indicator gene is a 15 reporter gene construct which including a reporter gene in operative linkage with one or more transcriptional regulatory elements responsive to the signal transduction activity of the target receptor, with the level of expression of the reporter gene providing the receptor-dependent detection signal. The amount of transcription from the reporter gene may be measured using any method known to those of skill in the art. For example, 20 specific mRNA expression may be detected using Northern blots or specific protein product may be identified by a characteristic stain or an intrinsic activity. In preferred embodiments, the gene product of the reporter is detected by an intrinsic activity associated with that product. For instance, the reporter gene may encode a gene product that, by enzymatic activity, gives rise to a detection signal based on color, fluorescence, 25 or luminescence.

The amount of activation of the indicator gene, e.g., expression of a reporter gene, is then compared to the amount of expression in either the same cell in the absence of the test compound or it may be compared with the amount of transcription in a substantially identical cell that lacks the specific receptors. A control cell may be 30 derived from the same cells from which the recombinant cell was prepared but which had not been modified by introduction of heterologous DNA, e.g., the encoding a test polypeptide. Alternatively, it may be a cell in which the specific receptors are removed. Any difference, e.g., a statistically significant difference, in the amount of transcription indicates that the test compound has in some manner altered the activity of the specific 35 receptor.

- 24 -

In other preferred embodiments, the reporter gene provides a selection method such that cells in which the compound is an effector for the receptor have a growth advantage. For example the reporter could enhance cell viability, relieve a cell nutritional requirement, and/or provide resistance to a drug.

5 By using any one of these readouts, compounds which modulate signaling via the heterologous receptor can be selected. If the compound does not appear to modulate signaling via the receptor protein, the assay may be repeated and modified by the introduction of a step in which the recombinant cell is first contacted with a known activator of the target receptor to induce signal transduction from the  
10 receptor, and the compound is assayed for its ability to inhibit the activity of the receptor, e.g., to identify receptor antagonists. In yet other embodiments, compounds can be screened for members which potentiate the response to a known activator of the receptor.

15 III. Host Cells

The host cells of the present invention may be of any species of yeast which are cultivable and in which an exogenous receptor can be made to engage the appropriate signal transduction machinery of the host cell. Exemplary species include Kluyveri lactis, Schizosaccharomyces pombe, and Ustilago maydis, with  
20 Saccharomyces cerevisiae being preferred. Other yeast which can be used in practicing the present invention are Neurospora crassa, Aspergillus niger, Aspergillus nidulans, Pichia pastoris, Candida tropicalis, and Hansenula polymorpha. The term "yeast", as used herein, includes not only yeast in a strictly taxonomic sense, i.e., unicellular organisms, but also yeast-like multicellular fungi or filamentous fungi.

25 The choice of appropriate host cell will also be influenced by the choice of detection signal. For instance, reporter constructs, as described below, can provide a selectable or screenable trait upon transcriptional activation (or inactivation) in response to a signal transduction pathway coupled to the target receptor. The indicator gene may be an unmodified gene already in the host cell pathway, such as the genes responsible  
30 for growth arrest in yeast. In certain embodiments a host cell gene may be operably linked to a "receptor-responsive" promoter. Alternatively, it may be a heterologous gene that has been so linked. Suitable genes and promoters are discussed below.

To achieve optimal selection or screening, the host cell phenotype will be considered. For example, introducing a pheromone-responsive chimeric HIS3 gene into  
35 a yeast that has a wild-type HIS3 gene would frustrate genetic selection. Thus, to achieve nutritional selection, an auxotrophic strain is preferred. Yeast strains that are

auxotrophic for histidine (HIS3) are known, see Struhl and Hill, (1987) Mol. Cell. Biol., 7:104; Fasullo and Davis, Mol. Cell. Biol., (1988) 8:4370. The HIS3 (imidazoleglycerol phosphate dehydratase) gene has been used as a selective marker in yeast. See Sikorski and Heiter, (1989) Genetics, 122:19; Struhl, et al., P.N.A.S. (1979) 76:1035; and, for 5 FUS1-HIS3 fusions, see Stevenson, et al., (1992) Genes Dev., 6:1293.

In certain embodiments, the host yeast cell can be modified in other ways. For example, it may be desirable to inactivate, such as by mutation or deletion, a homologous receptor, e.g., a pheromone receptor, present in the cell in order to minimize interference with signaling via the heterologous receptor. "Inactivation", with 10 respect to genes of the host cell, means that production of a functional gene product is prevented or inhibited. Inactivation may be achieved by deletion of the gene, mutation of the promoter so that expression does not occur, or mutation of the coding sequence so that the gene product is inactive. Inactivation may be partial or total.

15 In a preferred embodiment of the subject assay, the yeast cells possess one or more of the following characteristics: (a) the endogenous FUS1 gene has been inactivated; (b) the endogenous SST2 gene, and/or other genes involved in desensitization, have been inactivated; (c) if there is a homologous, endogenous receptor gene it has been inactivated; and (d) if the yeast produces an endogenous ligand to the exogenous receptor, the genes encoding for the ligand been inactivated.

20 It is desirable that the exogenous receptor be exposed on a continuing basis to the peptides. In some instances, this may result in desensitization of the pheromone pathway to the stimulus. For example, the mating signal transduction pathway is known to become desensitized by several mechanisms including pheromone degradation and modification of the function of the receptor, G proteins and/or downstream elements of the pheromone signal transduction by the products of the SST2, STE50, AFR1 (Konopka, J.B. (1993) Mol. Cell. Biol. 13:6876-6888) and SGV1, MSG5, and SIG1 genes. Selected mutations in these genes can lead to hypersensitivity to pheromone and an inability to adapt to the presence of pheromone. For example, introduction of mutations that interfere with function into strains expressing 25 heterologous G protein-coupled receptors constitutes a significant improvement on wild type strains and enables the development of extremely sensitive bioassays for compounds that interact with the receptors. Other mutations e.g. STE50, sgv1,bar1, ste2,ste3,pik1,msg5, sig1, and aft1, have the similar effect of increasing the sensitivity of the bioassay. Thus desensitization may be avoided by mutating (which may include 30 deleting) the SST2 gene so that it no longer produces a functional protein, or by mutating one of the other genes listed above.

- 26 -

In certain embodiments, it will be desirable to complement the host yeast cells, e.g., least partial function of an inactivated gene of the host cell can be supplied by an exogenous nucleic acid. For instance, yeast cells can be "mammalianized", and even "humanized", by complementation of receptor and signal transduction proteins with 5 mammalian homologues. To illustrate, inactivation of a yeast *Byr2/Stell1* gene can be complemented by expression of a human MEKK gene.

Complementations for use in the subject assay can be constructed without any undue experimentation. Indeed, many yeast genetic complementations with mammalian signal transduction proteins have been described in the art. For example, 10 Mosteller et al. (1994) Mol Cell Biol 14:1104-12 demonstrates that human Ras proteins can complement loss of ras mutations in *S. cerevisiae*. Moreover, Toda et al. (1986) Princess Takamatsu Symp 17: 253-60 have shown that human ras proteins can complement the loss of RAS1 and RAS2 proteins in yeast, and hence are functionally 15 homologous. Both human and yeast RAS proteins can stimulate the magnesium and guanine nucleotide-dependent adenylate cyclase activity present in yeast membranes. Ballester et al. (1989) Cell 59: 681-6 describe a vector to express the mammalian GAP protein in the yeast *S. cerevisiae*. When expressed in yeast, GAP inhibits the function of the human ras protein, and complements the loss of IRA1. IRA1 is a yeast gene that encodes a protein with homology to GAP and acts upstream of RAS. Mammalian GAP 20 can therefore function in yeast and interact with yeast RAS. Wei et al. (1994) Gene 151: 279-84 describes that a human Ras-specific guanine nucleotide-exchange factor, Cdc25GEF, can complement the loss of CDC25 function in *S. cerevisiae*. Martegani et al. (1992) EMBO J 11: 2151-7 describe the cloning by functional complementation of a mouse cDNA encoding a homolog of CDC25, a *Saccharomyces cerevisiae* RAS 25 activator. Vojtek et al. (1993) J Cell Sci 105: 777-85 and Matviw et al. (1992) Mol Cell Biol 12: 5033-40 describe how a mouse CAP protein, e.g., an adenylyl cyclase associated protein associated with ras-mediated signal transduction, can complements defects in *S. cerevisiae*. Papasavvas et al. (1992) Biochem Biophys Res Commun 184:1378-85 also suggest that inactivated yeast adenyl cyclase can be complemented by 30 a mammalian adenyl cyclase gene. Hughes et al. (1993) Nature 364: 349-52 describe the complementation of *byr1* in fission yeast by mammalian MAP kinase kinase (MEK). Parissenti et al. (1993) Mol Cell Endocrinol 98: 9-16 describes the reconstitution of bovine protein kinase C (PKC) in yeast. The Ca(2+) and phospholipid-dependent 35 Ser/Thr kinase PKC plays important roles in the transduction of cellular signals in mammalian cells. Marcus et al. (1995) PNAS 92: 6180-4 suggests the complementation

of shk1 null mutations in *S. pombe* by either the structurally related *S. cerevisiae* Ste20 or mammalian p65PAK protein kinases.

#### IV. Expression Systems

5 In general, it will be desirable that an expression vector be capable of replication in the host cell. Heterologous DNA may be integrated into the host genome, and thereafter is replicated as a part of the chromosomal DNA, or it may be DNA which replicates autonomously, as in the case of a plasmid. In the latter case, the vector will include an origin of replication which is functional in the host. In the case of an  
10 integrating vector, the vector may include sequences which facilitate integration, e.g., sequences homologous to host sequences, or encoding integrases.

Appropriate cloning and expression vectors for use with bacterial, fungal, yeast, and mammalian cellular hosts are known in the art, and are described in, for example, Powels et al. (*Cloning Vectors: A Laboratory Manual*, Elsevier, New York, 15 1985). Mammalian expression vectors may comprise non-transcribed elements such as an origin of replication, a suitable promoter and enhancer linked to the gene to be expressed, and other 5' or 3' flanking nontranscribed sequences, and 5' or 3' nontranslated sequences, such as necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, and transcriptional termination sequences.

20 A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach et al. (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p. 83, incorporated by reference herein).  
25 These vectors can replicate in *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin can be used. Suitable promoters for function in yeast include the promoters for metallothionein, 3-phosphoglycerate kinase (Hitzeman et al., *J. Biol. Chem.* 255, 2073 (1980) or other glycolytic enzymes (Hess et al., *J. Adv. Enzyme Res.* 7, 149 (1968); and Holland et al. *Biochemistry* 17, 4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phospho-fructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phospho-glucose isomerase, and glucokinase. Suitable vectors and promoters for use in  
30 yeast expression are further described in R. Hitzeman et al., EPO Publn. No. 73,657.  
35 Other promoters, which have the additional advantage of transcription controlled by

growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned metallothionein and glyceraldehyde-3-phosphate dehydrogenase, as well as enzymes responsible for maltose and galactose utilization. Finally, promoters that are active in only one of the two haploid mating types may be appropriate in certain circumstances. Among these haploid-specific promoters, the pheromone promoters MFα1 and MFα1 are of particular interest.

#### V. Receptors

Numerous different receptor types can be expressed in yeast cells for use in the instant invention.

The "heterologous receptors" of the present invention may be any G protein-coupled receptor which is exogenous to the cell which is to be genetically engineered for the purpose of the present invention. This receptor may be, for example, a plant or animal cell receptor. Screening for binding to plant cell receptors may be useful in the development of, e.g., herbicides. In the case of an animal receptor, it may be of invertebrate or vertebrate origin. An invertebrate receptor would, for example, facilitate development of insecticides. The expression of a receptor from a different species of yeast is also included within the term "heterologous" and could be used in the development of fungicides. The receptor may also be a vertebrate, more preferably a mammalian, still more preferably a human, receptor. The exogenous receptor is also preferably a seven transmembrane segment receptor.

Known ligands for GPCRs include: purines and nucleotides, such as adenosine, cAMP, ATP, UTP, ADP, melatonin and the like; biogenic amines (and related natural ligands), such as 5-hydroxytryptamine, acetylcholine, dopamine, adrenaline, adrenaline, adrenaline, histamine, noradrenaline, noradrenaline, noradrenaline, tyramine/octopamine and other related compounds; peptides such as adrenocorticotropic hormone (acth), melanocyte stimulating hormone (msh), melanocortins, neurotensin (nt), bombesin and related peptides, endothelins, cholecystokinin, gastrin, neurokinin b (nk3), invertebrate tachykinin-like peptides, substance k (nk2), substance p (nk1), neuropeptide y (npy), thyrotropin releasing-factor (trf), bradykinin, angiotensin ii, beta-endorphin, c5a anaphalatoxin, calcitonin, chemokines (also called intercrines), corticotrophic releasing factor (crf), dynorphin, endorphin, fmlp and other formylated peptides, follitropin (fsh), fungal mating pheromones, galanin, gastric inhibitory polypeptide receptor (gip), glucagon-like peptides (glps), glucagon, gonadotropin releasing hormone (gnrh), growth hormone

- 29 -

releasing hormone(ghrh), insect diuretic hormone, interleukin-8, leutropin (lh/hcg), met-enkephalin, opioid peptides, oxytocin, parathyroid hormone (pth) and pthrp, pituitary adenylyl cyclase activating peptide (pacap), secretin, somatostatin, thrombin, thyrotropin (tsh), vasoactive intestinal peptide (vip), vasopressin, vasotocin; eicosanoids such as ip-  
5 prostacyclin, pg-prostaglandins, tx-thromboxanes; retinal based compounds such as vertebrate 11-cis retinal, invertebrate 11-cis retinal and other related compounds; lipids and lipid-based compounds such as cannabinoids, anandamide, lysophosphatidic acid, platelet activating factor, leukotrienes and the like; excitatory amino acids and ions such as calcium ions and glutamate.

10 Preferred G protein coupled receptors include:  $\alpha$ 1A-adrenergic receptor,  $\alpha$ 1B-adrenergic receptor,  $\alpha$ 2-adrenergic receptor,  $\alpha$ 2B-adrenergic receptor,  $\beta$ 1-adrenergic receptor,  $\beta$ 2-adrenergic receptor,  $\beta$ 3-adrenergic receptor, m1 acetylcholine receptor (AChR), m2 AChR, m3 AChR, m4 AChR, m5 AChR, D1 dopamine receptor, D2 dopamine receptor, D3 dopamine receptor, D4 dopamine receptor, D5 dopamine  
15 receptor, A1 adenosine receptor, A2b adenosine receptor, 5-HT1a receptor, 5-HT1b receptor, 5HT1-like receptor, 5-HT1d receptor, 5HT1d-like receptor, 5HT1d beta receptor, substance K (neurokinin A) receptor, fMLP receptor, fMLP-like receptor, angiotensin II type 1 receptor, endothelin ETA receptor, endothelin ETB receptor, thrombin receptor, growth hormone-releasing hormone (GHRH) receptor, vasoactive  
20 intestinal peptide receptor, oxytocin receptor, somatostatin SSTR1 and SSTR2, SSTR3, cannabinoid receptor, follicle stimulating hormone (FSH) receptor, leutropin (LH/HCG) receptor, thyroid stimulating hormone (TSH) receptor, thromboxane A2 receptor, platelet-activating factor (PAF) receptor, C5a anaphylatoxin receptor, Interleukin 8 (IL-8) IL-8RA, IL-8RB, Delta Opioid receptor, Kappa Opioid receptor, mip-1/RANTES  
25 receptor, Rhodopsin, Red opsin, Green opsin, Blue opsin, metabotropic glutamate mGluR1-6, histamine H2 receptor, ATP receptor, neuropeptide Y receptor, amyloid protein precursor receptor, insulin-like growth factor II receptor, bradykinin receptor, gonadotropin-releasing hormone receptor, cholecystokinin receptor, melanocyte stimulating hormone receptor, antidiuretic hormone receptor, glucagon receptor, and  
30 adrenocorticotrophic hormone II receptor.

Other suitable receptors are known in the art. The term "receptor," as used herein, encompasses both naturally occurring and mutant receptors.

Many of these G protein-coupled receptors, like the yeast  $\alpha$ - and  $\alpha$ -factor receptors, contain seven hydrophobic amino acid-rich regions which are assumed to lie within the plasma membrane. Thus, for expression in yeast, the gene could be operably linked to a promoter functional in the cell to be engineered and to a signal sequence that  
35

also functions in the cell. For example, suitable promoters include Ste2, Ste3 and gal10. Optionally, the codons of the gene would be optimized for expression in yeast. See Hoekema et al.,(1987) Mol. Cell. Biol., 7:2914-24; Sharp, et al., (1986) 14:5125-43.

In some instances a foreign receptor which is expressed in yeast will 5 functionally integrate into the yeast membrane, and there interact with the endogenous yeast G protein. In other instances, either the receptor may be modified or a compatible G protein or a chimeric (i.e., part yeast/part mammalian) G protein subunit which can properly interact with the exogenous receptor G protein may be provided. The homology of STRs is discussed in Dohlman et al., Ann. Rev. Biochem., (1991) 10 60:653-88. When STRs are compared, a distinct spatial pattern of homology is discernible. The transmembrane domains are often the most similar, whereas the N- and C-terminal regions, and the cytoplasmic loop connecting transmembrane segments V and VI are more divergent. The functional significance of different STR regions has been studied by introducing point mutations (both substitutions and deletions) and by 15 constructing chimeras of different but related STRs. Synthetic peptides corresponding to individual segments have also been tested for activity. Affinity labeling has been used to identify ligand binding sites. Such information can be useful in creating mutations in GPCRs to enhance functionality.

If a naturally occurring exogenous GPCR cannot be made functional in 20 yeast, it may be mutated for this purpose. A comparison would be made of the amino acid sequences of the exogenous receptor and of the yeast receptors, and regions of high and low homology identified. Trial mutations would then be made to distinguish regions involved in ligand or G protein binding, from those necessary for functional integration in the membrane. The exogenous receptor would then be mutated in the latter region to 25 more closely resemble the yeast receptor, until functional integration was achieved. If this were insufficient to achieve functionality, mutations would next be made in the regions involved in G protein binding.

Preferably, the yeast genome is modified so that it is unable to produce 30 the yeast receptors which are homologous to the exogenous receptors in functional form in order to facilitate assay interpretation. For example, the endogenous G protein or G protein subunit is mutated generating, for example, a temperature sensitive mutant.

A common technique which has been used for cloning receptors is by 35 nucleic acid hybridization technology to identify receptors which are homologous to other, known receptors. For instance, the cloning of a receptor can be accomplished by the isolation and sequencing of the corresponding protein or the use of expression cloning techniques based on sequence homologies between these receptors. This

technology, since it does not require previous knowledge of the ligand for the receptor, has resulted in the cloning of a large number of "orphan receptors", which have no known ligand and often whose biological function is obscure.

- Many orphan receptors are GPCRs, but receptors of all types comprise  
5 this large family. Known orphan receptors include the nuclear receptors COUP-TF1/EAR3, COUP-TF2/ARP1, EAR-1, EAR-2, TR-2, PPAR1, HNF-4, ERR-1, ERR-2, NGFIB/Nur77, ELP/SF-1 and MPL (Parker et al, *supra*, and Power et al. (1992) *TIBS* 13:318-323). A large number of orphan receptors have been identified in the EPH family (Hirai et al (1987) *Science* 238:1717-1720). HER3 and HER4 are orphan  
10 receptors in the epidermal growth factor receptor family (Plowman et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:1746-1750). ILA is a newly identified member of the human nerve growth factor/tumor necrosis factor receptor family (Schwarz et al. (1993) *Gene* 134:295-298). IRRR is an orphan insulin receptor-related receptor which is a transmembrane tyrosine kinase (Shier et al. (1989) *J. Biol Chem* 264:14606-14608).  
15 Several orphan tyrosine kinase receptors have been found in Drosophila (Perrimon (1994) *Curr. Opin. Cell Biol.* 6:260-266).

- In one aspect, the present invention provides for novel assays which have been optimized for testing for agonists and antagonists of GPCRs, in particular orphan GPCRs. These assays involve the use of four different strains of yeast, each of which  
20 has been engineered to contain a different complement of G protein subunits. This expression of various combinations of G protein subunits increases the likelihood of obtaining functional coupling even absent knowing which G protein subunit chimera (or combination of subunits) will best function in the screening assay. The importance of identifying ligands for orphan receptors is clear; it opens up a wide area for research in  
25 the area of drug discovery.

#### VI. G protein subunits and complexes

- In certain instances it will be desirable to modify naturally occurring forms of yeast or mammalian G-protein subunits. For instance, where a heterologous  
30 GPCR does not adequately couple to the endogenous yeast G protein subunit, such a subunit, e.g., GPA1 may be modified to improve coupling. Such modifications can be made by mutation, e.g., directed mutation or random mutation, using methods known in the art and described in more detail below.

- Alternatively, a heterologous subunit can be expressed. The specificity  
35 of coupling of a receptor to a heterotrimeric G-protein is largely determined by the  $\alpha$  subunit of the G-protein. Thus, in preferred embodiments, a heterologous  $G\alpha$  subunit is

expressed in the yeast cell. The predominant role of the yeast G $\alpha$ , GPA1, is to bind to and sequester the effector-signaling  $\beta\gamma$  component of the heterotrimer. Thus, in order to achieve functional integration into a yeast pheromone signaling pathway, a heterologous G $\alpha$  subunit must bind to yeast  $\beta\gamma$  in the quiescent state, and release it upon receptor activation.

If functional integration is not achieved, or is not optimal, the heterologous subunit can be mutated. For example, in general, mammalian G $\alpha$  subunits couple poorly to the  $\beta\gamma$  subunits of yeast cells. In yeast which lack their own endogenous G $\alpha$  subunit, this failure to couple results in the constitutive activation of the pheromone pathway due to the effector activity of the unbound yeast  $\beta\gamma$ . Accordingly, if a naturally occurring heterologous G protein subunit does not enhance coupling, modifications can be made. Such modifications may take the form of mutations which are designed to increase the resemblance of the G protein subunit to the yeast G protein subunit while decreasing its resemblance to the heterologous receptor-associated G protein subunit.

For example, a residue may be changed so as to become identical to the corresponding yeast G protein residue, or to belong to the same exchange group of that residue. After modification, the modified G protein subunit might or might not be "substantially homologous" to the heterologous and/or the yeast G protein subunit.

In the case of G $\alpha$ , modifications are preferably concentrated in regions of the G $\alpha$  which are likely to be involved in G $\beta\gamma$  binding.

In other embodiments, modifications will take the form of replacing one or more amino acids of the receptor-associated G protein subunit with the corresponding yeast G protein subunit amino acids, thereby forming a chimeric G protein subunit. In preferred embodiments, three or more consecutive amino acids are replaced. In other embodiments, point mutations may be sufficient.

Chimeric G protein subunits of the invention enhance coupling of the heterologous receptor to the endogenous yeast signaling pathway. For example, a chimeric G $\alpha$  subunit will interact with the heterologous receptor and the yeast G $\beta\gamma$  complex, thereby permitting signal transduction.

A yeast cell of the present invention can express one or more of the indicated G protein structures. For example, a yeast cell can express a chimeric G $\alpha$  subunit, and an endogenous yeast G $\beta\gamma$ , a mammalian G $\beta\gamma$ , a mutated mammalian G $\beta\gamma$ , or a chimeric G $\beta\gamma$ .

In preferred embodiments, both the receptor and the heterologous subunit are derived from the same source, e.g., are mammalian. In particularly preferred embodiment, both are human in origin.

5 In another preferred embodiment, a yeast cell that expresses a heterologous or chimeric G protein subunit has been modified such that the endogenous, homologous G protein subunit gene is disrupted.

In certain embodiments, yeast strains lacking pheromone receptors and having no heterologous receptor capable of coupling to the pheromone response pathway may be used to assess the affinity of an endogenous yeast G protein subunit, a mutated G protein subunit, or chimeric G protein subunit for other yeast subunits. For example, the affinity of *gpa1p*, chimeric *gpa-G $\alpha$ s*, or other  $G\alpha$  subunit for yeast  $\beta\gamma$  or other chimeric  $\beta\gamma$  subunit can be assessed. Such strains depend on free  $\beta\gamma$  for signaling through the pheromone response pathway leading to growth arrest. Mutant  $G\alpha$  subunits may be tested in such a system, those which bind  $\beta\gamma$  more effectively will sequester  $\beta\gamma$  and reduce or block signaling. Preferably, such chimeras and *gpa1* subunits can be assessed in a *gpa1<sup>-</sup>* background to avoid competition with *gpa1* for  $\beta\gamma$ . For example,  $G\alpha$ s chimeric mutants (see below) carrying D229S, E10K, N254D, or S286P were found to sequester  $\beta\gamma$  more effectively than the chimeras with wild type sequences. Also, double mutants were even more effective than either single mutant. Similarly, overexpression of  $G\alpha$ s by driving transcription from the highly efficient PGK promoter resulted in dampening of the receptor coupling which may be offset by introduction of the double mutant  $G\alpha$ s (D229S,E10K).

25 Guidance for making mutations in G protein subunits and in the construction of chimeric G protein subunits is provided below.

Site-Directed Mutagenesis versus Random Mutagenesis

There are numerous art recognized ways to solve the structure-function problems of the sort presented by attempts to define the determinants involved in mediating the association of the subunits that comprise the G protein heterotrimer. For example, in one approach, discussed above with respect to hybrid constructs, specific mutations or alterations are introduced into a molecule based upon the available experimental evidence. In a second approach, random mutagenesis techniques, coupled with selection or screening systems, are used to introduce large numbers of mutations into a molecule, and that collection of randomly mutated molecules is then subjected to a selection for the desired phenotype or a screen in which the desired phenotype can be observed against a background of undesirable phenotypes.

With random mutagenesis one can mutagenize an entire molecule or one can proceed by cassette mutagenesis. In the former instance, the entire coding region of a molecule is mutagenized by one of several methods (chemical, PCR, doped oligonucleotide synthesis) and that collection of randomly mutated molecules is 5 subjected to selection or screening procedures. Random mutagenesis can be applied in this way in cases where the molecule being studied is relatively small and there are powerful and stringent selections or screens available to discriminate between the different classes of mutant phenotypes that will inevitably arise. In the second approach, discrete regions of a protein, corresponding either to defined structural (i.e.  $\alpha$ -helices,  $\beta$  10 -sheets, turns, surface loops) or functional determinants (e.g.. catalytic clefts, binding determinants, transmembrane segments) are subjected to saturating or semi-random mutagenesis and these mutagenized cassettes are re-introduced into the context of the otherwise wild type allele.

Cassette mutagenesis is most useful when there is experimental evidence 15 available to suggest a particular function for a region of a molecule and there is a powerful selection and/or screening approach available to discriminate between interesting and uninteresting mutants. Cassette mutagenesis is also useful when the parent molecule is comparatively large and the desire is to map the functional domains of a molecule by mutagenizing the molecule in a step-wise fashion, i.e. mutating one 20 linear cassette of residues at a time and then assaying for function.

The present invention provides for applying random mutagenesis in order to further delineate the determinants involved in  $G\alpha$ - $G\beta\gamma$  or subunit-receptor association. Random mutagenesis may be accomplished by many means, including:

1. PCR mutagenesis, in which the error prone Taq polymerase is 25 exploited to generate mutant alleles of G protein subunits, which are assayed directly in yeast for an ability to couple.
  2. Chemical mutagenesis, in which expression cassettes encoding G protein subunits are exposed to mutagens and the protein products of the mutant sequences are assayed directly in yeast for an ability to couple.
  3. Doped synthesis of oligonucleotides encoding portions of the G protein 30 subunit gene.
  4. In vivo mutagenesis, in which random mutations are introduced into the coding region of G protein subunits by passage through a mutator strain of *E. coli*, XL1-Red (*mutD5 mutS mutT*) (Stratagene, Menasha, WI).
- 35 In certain embodiments, for example, the random mutagenesis may be focused on regions suspected to be involved in  $G\alpha$ - $G\beta\gamma$  association. Random

mutagenesis approaches are feasible for two reasons. First, in yeast one has the ability to construct stringent screens and facile selections (growth vs. death, transcription vs. lack of transcription) that are not readily available in mammalian systems. Second, when using yeast it is possible to screen efficiently through thousands of transformants  
5 rapidly. For example, this relatively small region of G $\alpha$  subunits represents a reasonable target for cassette mutagenesis. Another region that may be amenable to cassette mutagenesis is that defining the surface of the switch region of G $\alpha$  subunits that is solvent-exposed in the crystal structures of G $\alpha$ i and transducin. From the data described below, this surface may contain residues that are in direct contact with yeast G $\beta$  $\gamma$   
10 subunits, and may therefore be a reasonable target for mutagenesis.

#### A. Modification of G $\alpha$

Some aspects of G $\alpha$  structure are relevant to the design of modified G $\alpha$  subunits. Alignments of G $\alpha$  and GPA1 can be made to determine sequence similarity.  
15 For alignments of the entire coding regions of GPA1 with G $\alpha$ s, G $\alpha$ i, and G $\alpha$ O, G $\alpha$ q and G $\alpha$ z, see Dietzel and Kurjan (1987, Cell 50:573) and Lambright, et al. (1994, Nature 369:621-628). Additional sequence information is provided by Mattera, et al. (1986, FEBS Lett 206:36-41), Bray, et al. (1986, Proc. Natl. Acad. Sci USA 83:8893-8897) and Bray, et al. (1987, Proc Natl. Acad Sci USA 84:5115-5119). An alignment of GPA1 and  
20 four other G $\alpha$  proteins is provided by Stone and Reed (1990, Mol. Cell Biol. 10:4439). See also the alignment presented in Figure 1.

The gene encoding a G protein homolog of *S. cerevisiae* was cloned independently by Dietzel and Kurjan (*supra*) (who referred to the gene as SCG1) and by Nakafuku, et al. (1987 Proc Natl Acad Sci 84:2140-2144) (who called the gene GPA1).  
25 Sequence analysis revealed a high degree of homology between the protein encoded by this gene and mammalian G $\alpha$ . GPA1 encodes a protein of 472 amino acids, as compared with approximately 340-350 amino acids for most mammalian G $\alpha$  subunits in four described families, G $\alpha$ s, G $\alpha$ i, G $\alpha$ q and G $\alpha$ 12/13. Nevertheless, GPA1 shares overall sequence and structural homology with all G $\alpha$  proteins identified to date. The highest  
30 overall homology in GPA1 is to the G $\alpha$ i family (48% identity, or 65% with conservative substitutions) and the lowest is to G $\alpha$ s (33% identity, or 51% with conservative substitutions) (Nakafuku, et al., *supra*).

The regions of high sequence homology among G $\alpha$  subunits are dispersed throughout their primary sequences, with the regions sharing the highest  
35 degree of homology mapping to sequence that comprises the guanine nucleotide binding/GTPase domain. This domain is structurally similar to the  $\alpha\beta$  fold of ras

proteins and the protein synthesis elongation factor EF-Tu. This highly conserved guanine nucleotide-binding domain consists of a six-stranded  $\beta$  sheet surrounded by a set of five  $\alpha$ -helices. It is within these  $\beta$  sheets and  $\alpha$  helices that the highest degree of conservation is observed among all G $\alpha$  proteins, including GPA1. The least sequence and structural homology is found in the intervening loops between the  $\beta$  sheets and  $\alpha$  helices that define the core GTPase domain. There are a total of four "intervening loops" or "inserts" present in all G $\alpha$  subunits. In the crystal structures reported to date for the GDP- and GTP $\gamma$ S-ligated forms of bovine rod transducin (Noel, et al. (1993) Nature 366:654-663); (Lambright, et al. (1994) Nature 369:621-628), the loop residues are found to be outside the core GTPase structure. Functional roles for these loop structures have been established in only a few instances. A direct role in coupling to phosphodiesterase- $\gamma$  has been demonstrated for residues within inserts 3 and 4 of G $\alpha$ i (Rarick, et al. (1992) Science 256:1031-1033); (Artemyev, et al. (1992) J. Biol. Chem. 267:25067-25072), while a "GAP-like" activity has been ascribed to the largely  $\alpha$ -helical insert 1 domain of G $\alpha$ S (Markby, et al. (1993) Science 262:1805-1901).

While the amino- and carboxy-termini of G $\alpha$  subunits do not share striking homology either at the primary, secondary, or tertiary levels, there are several generalizations that can be made about them. First, the amino termini of G $\alpha$  subunits have been implicated in the association of G $\alpha$  with G $\beta$  $\gamma$  complexes and in membrane association via N-terminal myristylation. In addition, the carboxy-termini have been implicated in the association of G $\alpha$  $\beta$  $\gamma$  heterotrimeric complexes with G protein-coupled receptors (Sullivan, et al. (1987) Nature 330:758-760); West, et al. (1985) J. Biol. Chem. 260:14428-14430); (Conklin, et al. (1993) Nature 363:274-276); (Kallal and Kurjan. 1997. Mol. Cell. Biol. 17:2897). Data in support of these generalizations about the function of the N-terminus derive from several sources, including both biochemical and genetic studies.

Figure 1 shows the amino terminal 66 residues of GPA1 aligned with the cognate domains of human G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3, G $\alpha$ 16 and transducin. In the GPA41G $\alpha$  hybrids, the amino terminal 41 residues (derived from GPA1) are identical and end with the sequence-LEKQRDKNE-(SEQ ID NO:79). All residues following the glutamate (E) residue at position 41 are contributed by the human G $\alpha$  subunits, including the consensus nucleotide binding motif shown in the amino acid sequence -GxGxxG-(SEQ ID NO:80). Periods in the sequences indicate gaps that have been introduced to maximize alignments in this region. Codon bias is mammalian. For alignments of the entire coding regions of GPA1 with G $\alpha$ s, G $\alpha$ i, and G $\alpha$ O, G $\alpha$ q and G $\alpha$ z, see Dietzel and Kurjan (1987, Cell 50:573) and Lambright, et al. (1994, Nature 369:621-628).

Additional sequence information is provided by Mattera, *et al.* (1986, FEBS Lett 206:36-41), Bray, *et al.* (1986, Proc. Natl. Acad. Sci USA 83:8893-8897) and Bray, *et al.* (1987, Proc Natl. Acad Sci USA 84:5115-5119).

As indicated above, there is little if any sequence homology shared among the amino termini of G $\alpha$  subunits. The amino terminal domains of G $\alpha$  subunits that precede the first  $\beta$ -sheet (containing the sequence motif -LLLLGAGESG-(SEQ ID NO:81); see Noel, *et al.* (*supra*) for the numbering of the structural elements of G $\alpha$  subunits) vary in length from 41 amino acids (GPA1) to 31 amino acids (G $\alpha$ t). Most G $\alpha$  subunits share the consensus sequence for the addition of myristic acid at their amino termini (MGXaaS-) (SEQ ID NO:82), although not all G $\alpha$  subunits that contain this motif have myristic acid covalently associated with the glycine at position 2 (Speigel, *et al.* (1991) TIBS 16:338-3441). The role of this post-translational modification has been inferred from studies in which the activity of mutant G $\alpha$  subunits from which the consensus sequence for myristoylation has been added or deleted has been assayed (Mumby *et al.* (1990) Proc. Natl. Acad. Sci. USA 87: 728-732; (Linder, *et al.* (1991) J. Biol Chem. 266:4654-4659); Gallego, *et al.* (1992) Proc. Natl. Acad. Sci. USA 89:9695-9699). These studies suggest two roles for N-terminal myristoylation. First, the presence of amino-terminal myristic acid has in some cases been shown to be required for association of G $\alpha$  subunits with the membrane, and second, this modification has been demonstrated to play a role in modulating the association of G $\alpha$  subunits with G $\beta\gamma$  complexes. The role of myristoylation of the GPA1 gene products, at present is unknown.

In other biochemical studies aimed at examining the role of the amino-terminus of G $\alpha$  in driving the association between G $\alpha$  and G $\beta\gamma$  subunits, proteolytically or genetically truncated versions of G $\alpha$  subunits were assayed for their ability to associate with G $\beta\gamma$  complexes, bind guanine nucleotides and/or to activate effector molecules. In all cases, G $\alpha$  subunits with truncated amino termini were deficient in all three functions (Graf, *et al.* (1992) J. Biol. Chem. 267:24307-24314); (Journot, *et al.* (1990) J. Biol. Chem. 265:9009-9015); and (Neer, *et al.* (1988) J. Biol. Chem. 263:8996-9000). Slepak, *et al.* (1993, J. Biol. Chem. 268:1414-1423) reported a mutational analysis of the N-terminal 56 amino acids of mammalian G $\alpha$ o expressed in Escherichia coli. Molecules with an apparent reduced ability to interact with exogenously added mammalian G $\beta\gamma$  were identified in the mutant library. As the authors pointed out, however, the assay used to screen the mutants the extent of ADP-ribosylation of the mutant G $\alpha$  by pertussis toxin was not a completely satisfactory probe of interactions between G $\alpha$  and G $\beta\gamma$ . Mutations identified as inhibiting the

interaction of the subunits, using this assay, may still permit the complexing of G $\alpha$  and G $\beta\gamma$  while sterically hindering the ribosylation of G $\alpha$  by toxin. Other work has revealed specific amino acid residues of GPA1 that are important in GPA1 function. For example, a E307K mutation appears to create an a subunit with a broadened specificity for G $\beta$  subunits (Whiteway *et al.* 1994. Mol. Cell. Biol. 14:3223). Interestingly, the residue in the mammalian G  $\alpha$  subunit which is equivalent to the E307 position is diagnostic for a particular class of mammalian  $\alpha$  subunits. For example, the G<sub>s</sub> $\alpha$  subunits contain a lysine at this position, the G<sub>o</sub> and G<sub>i</sub>  $\alpha$  subunits contain a histidine, the transducin  $\alpha$  subunits have a glutamine, the Gq  $\alpha$  subunits have a proline, and the G<sub>13</sub>  $\alpha$  subunits have an aspartic acid at this site (Whiteway *et al.* supra).

Genetic studies examined the role of amino-terminal determinants of G $\alpha$  in heterotrimer subunit association have been carried out in both yeast systems using GPA1-mammalian G $\alpha$  hybrids (Kang, *et al.* (1990) Mol. Cell. Biol. 10:2582-2590) and in mammalian systems using G $\alpha$ i/G $\alpha$ s hybrids (Russell and Johnson (1993) Mol. Pharmacol. 44:255-263). In the former studies, gene fusions, composed of yeast GPA1 and mammalian G $\alpha$  sequences were constructed by Kang, *et al.* (supra) and assayed for their ability to complement a gpal null phenotype (i.e., constitutive activation of the pheromone response pathway) in *S. cerevisiae*. Kang, *et al.* demonstrated that wild type mammalian G $\alpha$ s, G $\alpha$ i but not G $\alpha$ o proteins are competent to associate with yeast G $\alpha$  and suppress the gpal null phenotype, but only when overexpressed. Fusion proteins containing the amino-terminal 330 residues of GPA1 sequence linked to 160, 143, or 142 residues of the mammalian G $\alpha$ s, G $\alpha$ i and G $\alpha$ o carboxyl-terminal regions, respectively, also coupled to the yeast mating response pathway when overexpressed on high copy plasmids with strong inducible (CUP) or constitutive (PGK) promoters. All three of these hybrid molecules were able to complement the gpal null mutation in a growth arrest assay, and were additionally able to inhibit  $\alpha$ -factor responsiveness and mating in tester strains. These last two observations argue that hybrid yeast-mammalian G $\alpha$  subunits are capable of interacting directly with yeast G $\beta\gamma$ , thereby disrupting the normal function of the yeast heterotrimer. Fusions containing the amino terminal domain of G $\alpha$ s, G $\alpha$ i or G $\alpha$ o, however, did not complement the gpal null phenotype, indicating a requirement for determinants in the amino terminal 330 amino acid residues of GPA1 for association and sequestration of yeast G $\beta\gamma$  complexes. Taken together, these data suggest that determinants in the amino terminal region of G $\alpha$  subunits determine not only the ability to associate with G $\beta\gamma$  subunits in general, but also with specific G $\beta\gamma$  subunits in a species-restricted manner.

Hybrid G $\alpha$ i/G $\alpha$ s subunits have been assayed in mammalian expression systems (Russell and Johnson (*supra*). In these studies, a large number of chimeric G $\alpha$  subunits were assayed for an ability to activate adenylyl cyclase, and therefore, indirectly, for an ability to interact with G $\beta$  $\gamma$  (i.e., coupling of G $\alpha$  to G $\beta$  $\gamma$ = inactive cyclase; uncoupling of G $\alpha$  from G $\beta$  $\gamma$  = active cyclase). From these studies a complex picture emerged in which determinants in the region between residues 25 and 96 of the hybrids were found to determine the state of activation of these alleles as reflected in their rates of guanine nucleotide exchange and GTP hydrolysis and the extent to which they activated adenylyl cyclase *in vivo*. These data could be interpreted to support the hypothesis that structural elements in the region between the amino terminal methionine and the  $\beta$  sheet identified in the crystal structure of G $\alpha$ t (see Noel, et al. *supra* and Lambright, et al. *supra*) are involved in determining the state of activity of the heterotrimer by (1) driving association/dissociation between G $\alpha$  and G $\beta$  $\gamma$  subunits; (2) driving GDP/GTP exchange. While there is no direct evidence provided by these studies to support the idea that residues in this region of G $\alpha$  and residues in G $\beta$  $\gamma$  subunits contact one another, the data nonetheless provide a positive indication for the construction of hybrid G $\alpha$  subunits that retain function. There is, however, a negative indicator that derives from this work in that some hybrid constructs resulted in constitutive activation of the chimeric proteins (i.e., a loss of receptor-dependent stimulation of G $\beta$  $\gamma$  dissociation and effector activation).

#### B. Construction of chimeric G $\alpha$ subunits.

In preferred embodiments chimeric G $\alpha$  subunits retain as much of the sequence of the native mammalian proteins as possible and, in particularly preferred embodiments, the level of expression for the heterologous components should approach, as closely as possible, the level of their endogenous counterparts. The results described by King, et al. (1990, *Science* 250:121-123) for expression of the human  $\beta$ 2-adrenergic receptor and G $\alpha$ s in yeast, taken together with negative results obtained by Kang, et al. (*supra*) with full-length mammalian G $\alpha$  subunits other than G $\alpha$ s, led to the following preferred embodiments for the development of yeast strains in which mammalian G protein-coupled receptors could be linked to the pheromone response pathway.

In one embodiment, mammalian G $\alpha$  subunits are expressed using the native sequence of each subunit or, alternatively, as minimal gene fusions with sequences from the amino- terminus of GPA1 replacing the homologous residues from the mammalian G $\alpha$  subunits. In another embodiment, mammalian G $\alpha$  subunits are expressed from the GPA1 promoter either on low copy plasmids or after integration into

the yeast genome as a single copy gene. In certain embodiments, endogenous G $\beta\gamma$  subunits are provided by the yeast STE4 and STE18 loci, while in other embodiments chimeric or heterologous G $\beta$  and/or G $\gamma$  subunits are also provided.

5

### C. Rational Design of Chimeric G $\alpha$ Subunits

Several classes of rationally designed GPAI-mammalian G $\alpha$  hybrid subunits have been tested for the ability to couple to yeast G $\beta\gamma$ . The first, and largest, class of hybrids are those that encode different lengths of the GPAI amino terminal domain in place of the homologous regions of the mammalian G $\alpha$  subunits. This class of hybrid molecules includes GPA<sub>BAMH1</sub>, GPA<sub>41</sub>, GPA<sub>ID</sub>, and GPA<sub>LW</sub> hybrids, described below. The rationale for constructing these hybrid G $\alpha$  proteins is based on results, described above, that bear on the importance of the amino terminal residues of G $\alpha$  in mediating interaction with G $\beta\gamma$ .

Preferably, the yeast G $\alpha$  subunit is replaced by a chimeric G $\alpha$  subunit in which a portion, e.g., at least about 20, more preferably at least about 40, amino acids, from the amino terminus of the yeast G $\alpha$ , is fused to a sequence from a mammalian (or other exogenous) G $\alpha$ . While about 40 amino acids is the suggested starting point, shorter or longer portions may be tested to determine the minimum length required for coupling to yeast G $\beta\gamma$  and the maximum length compatible with retention of coupling to the exogenous receptor. It is presently believed that only the final 10 or 20 amino acids at the carboxy terminus of the G $\alpha$  subunit are required for interaction with the receptor.

#### i. GPA<sub>BAMH1</sub> hybrids.

Kang et al. supra. described hybrid G $\alpha$  subunits encoding the amino terminal 310 residues of GPAI fused to the carboxyl terminal 160, 143 and 142 residues, respectively, of G $\alpha$ S, G $\alpha$ I2, and G $\alpha$ O. In all cases examined by Kang et al., the hybrid proteins were able to complement the growth arrest phenotype of gpal strains. Hybrids between GPAI and G $\alpha$ I3, G $\alpha$ q and G $\alpha$ 16 have been constructed and functionally complement the growth arrest phenotype of gpal strains.

30

GPA<sub>41</sub> hybrids. The rationale for constructing a minimal hybrid encoding only 41 amino acids of GPAI relies upon the biochemical evidence for the role of the amino-terminus of G $\alpha$  subunits discussed above, together with the following observation. G $\beta$  and G $\gamma$  subunits are known to interact via  $\alpha$ -helical domains at their respective amino-termini (Pronin, et al. (1992) Proc. Natl. Acad. Sci. USA 89:6220-6224); Garritsen, et al. 1993). The suggestion that the amino termini of G $\alpha$  subunits may form an helical coil and that this helical coil may be involved in association of G $\alpha$  with

- 41 -

G $\beta$  $\gamma$  (Masters et al (1986) Protein Engineering 1:47-54); Lupas et al.(1992) FEBS Lett. 314:105-108) leads to the hypothesis that the three subunits of the G-protein heterotrimer interact with one another reversibly through the winding and unwinding of their amino-terminal helical regions. A mechanism of this type has been suggested, as well, from an analysis of leucine zipper mutants of the GCN4 transcription factor (Harbury, et al. (1993) Science 262:1401-1407). The rationale for constructing hybrids like those described by Kang, et al. *supra.*, that contain a majority of yeast sequence and only minimal mammalian sequence, derives from their ability to function in assays of coupling between G $\alpha$  and G $\beta$  $\gamma$  subunits. However, these chimeras had never been assayed for an ability to couple to both mammalian G protein-coupled receptors and yeast G $\beta$  $\gamma$  subunits, and hence to reconstitute a hybrid signaling pathway in yeast.

GPA41 hybrids that have been constructed and tested include G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3, G $\alpha$ q, G $\alpha$ o $a$ , G $\alpha$ o $b$  and G $\alpha$ 16. Hybrids of G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3, and G $\alpha$ 16 functionally complement the growth arrest phenotype of gpa1 strains, while GPA41 hybrids of G $\alpha$ o $a$  and G $\alpha$ o $b$  do not. In addition to being tested in a growth arrest assay, these constructs have been assayed in the more sensitive transcriptional assay for activation of a fuslp-HIS3 gene. In both of these assays, the GPA41-G $\alpha$ s hybrid couples less well than the GPA41-i2, -i3, and -16 hybrids, while the GPA41 -o $a$ , and -o $b$  hybrids do not function in either assay.

Several predictive algorithms indicate that the amino terminal domain up to the highly conserved sequence motif-LLLLGAGESG- (SEQ ID NO: 1) (the first L in this motif is residue 43 in GPA1) forms a helical structure with amphipathic character. Assuming that a heptahelical repeat unit, the following hybrids between yeast GPA1 and mammalian G $\alpha$ S can be used to define the number of helical repeats in this motif necessary for hybrid function:

GPA1-7/G $\alpha$ s8-394  
GPA1-14/G $\alpha$ s15-394  
GPA1-21/G $\alpha$ s22-394  
GPA1-28/G $\alpha$ s29-394  
30 GPA1-35/G $\alpha$ s36-394  
GPA1-42/G $\alpha$ s43-394

In these hybrids, the prediction is that the structural repeat unit in the amino terminal domain up to the tetra-leucine motif is 7, and that swapping sequences in units of 7 will in effect amount to a swap of unit turns of turns of the helical structure that comprises this domain.

- 42 -

A second group of "double crossover" hybrids of this class are those that are aligned on the first putative heptad repeat beginning with residue G11 in GPA1. In these hybrids, helical repeats are swapped from GPA1 into a G $\alpha$ S backbone one heptad repeat unit at a time.

- 5                   G $\alpha$ S1-10/GPA11-17/G $\alpha$ s18-394  
                  G $\alpha$ S1-17/GPA18-24/G $\alpha$ S25-394  
                  G $\alpha$ S1-17/GPA25-31/G $\alpha$ S32-394  
                  G $\alpha$ S1-17/GPA32-38/G $\alpha$ S39-394

The gap that is introduced between residues 9 and 10 in the G $\alpha$ S sequence is to preserve the alignment of the -LLLLGAGE-(positions 1-8 of SEQ ID NO:1)sequence motif. This class of hybrids can be complemented by cassette mutagenesis of each heptad repeat followed by screening of these collections of "heptad" libraries in standard coupling assays.

A third class of hybrids based on the prediction that the amino terminus forms a helical domain with a heptahelical repeat unit are those that effect the overall hydrophobic or hydrophilic character of the opposing sides of the predicted helical structure (See Lupas et al. supra). In this model, the a and d positions of the heptad repeat abcdefg are found to be conserved hydrophobic residues that define one face of the helix, while the e and g positions define the charged face of the helix. In this class of hybrids, the sequence of the G $\alpha$ S parent is maintained except for specific substitutions at one or more of the following critical residues to render the different helical faces of G $\alpha$ S more "GPA1-like"

- 25                   K8Q  
                  +I-10  
                  E10G  
                  Q12E  
                  R13S  
                  N14D  
30                   E15P  
                  E15F  
                  K17L  
                  E21R  
                  K28Q  
35                   K32L  
                  V36R

This collection of single mutations could be screened for coupling efficiency to yeast G $\beta\gamma$  and then constructed in combinations (double and greater if necessary).

A fourth class of hybrid molecules that span this region of GPA1-G $\alpha$  hybrids are those that have junctions between GPA1 and G $\alpha$  subunits introduced by three primer PCR. In this approach, the two outside primers are encoded by sequences at the initiator methionine of GPA1 on the 5' side and at the tetraleucine motif of G $\alpha$ S (for example) on the 3' side. A series of junctional primers spanning different junctional points can be mixed with the outside primers to make a series of molecules each with different amounts of GPA1 and G $\alpha$ S sequences, respectively.

#### ii. GPA<sub>ID</sub> and GPA<sub>LW</sub> hybrids.

The regions of high homology among G $\beta\gamma$  subunits that have been identified by sequence alignment are interspersed throughout the molecule. The G1 region containing the highly conserved -GSGESGDST- (SEQ ID NO: 2) motif is followed immediately by a region of very low sequence conservation, the "il" or insert 1 region. Both sequence and length vary considerably among the il regions of the G $\alpha$  subunits. By aligning the sequences of G $\alpha$  subunits, the conserved regions bounding the il region were identified and two additional classes of GPA1-G $\alpha$  hybrids were constructed. The GPA<sub>ID</sub> hybrids encode the amino terminal 102 residues of GPA1 (up to the sequence -QARKLGIQ-) (SEQ ID NO: 3) fused in frame to mammalian G $\alpha$  subunits, while the GPA<sub>LW</sub> hybrids encode the amino terminal 244 residues of GPA1 (up to the sequence LIHEDIAKA- (SEQ ID NO: 4) in GPA1). The reason for constructing the GPA<sub>ID</sub> and GPA<sub>LW</sub> hybrids was to test the hypothesis that the il region of GPA1 is required for mediating the interaction of GPA1 with yeast G $\beta\gamma$  subunits, for the stable expression of the hybrid molecules, or for function of the hybrid molecules. The GPA<sub>ID</sub> hybrids contain the amino terminal domain of GPA1 fused to the il domain of mammalian subunits, and therefore do not contain the GPA1 il region, while the GPA<sub>LW</sub> hybrids contain the amino terminal 244 residues of GPA1 including the entire il region (as defined by sequence alignments). Hybrids of both GPA<sub>ID</sub> and GPA<sub>LW</sub> classes were constructed for G $\alpha$ S, C- $\alpha$ i2, G $\alpha$ i3, G $\alpha$ o<sub>a</sub>, and G $\alpha$ 16; none of these hybrids complemented the gpal growth arrest phenotype.

Subsequent to the construction and testing of the GPA<sub>ID</sub> and GPA<sub>LW</sub> classes of hybrids, the crystal structures of G<sub>transducin</sub> in both the GDP and GTP $\gamma$ S-ligated form, and the crystal structure of several G $\alpha$ il variants in the GTP $\gamma$ S-ligated and GDP-AlF<sub>4</sub> forms were reported (Noel et al. supra; Lambright et al.

- 44 -

supra; and Coleman et al.(1994) Science 265:1405-1412). The crystal structures reveal that the il region defined by sequence alignment has a conserved structure that is comprised of six alpha helices in a rigid array, and that the junctions chosen for the construction of the GPA<sub>ID</sub> and GPA<sub>LW</sub> hybrids were not compatible with conservation 5 of the structural features of the il region observed in the crystals. The junction chosen for the GPA<sub>ID</sub> hybrids falls in the center of the long  $\alpha$ A helix; chimerization of this helix in all likelihood destabilizes it and the protein structure in general. The same is true of the junction chosen for the GPA<sub>LW</sub> hybrids in which the crossover point between GPA1 and the mammalian G $\alpha$  subunit falls at the end of the short  $\alpha$ C helix and therefore may 10 distort it and destabilize the protein.

The failure of the GPA<sub>ID</sub> and GPA<sub>LW</sub> hybrids is predicted to be due to disruption of critical structural elements in the il region as discussed above. Based upon new alignments and the data presented in Noel et al (supra), Lambright et al (supra), and Coleman et al (supra), this problem can be averted with the ras-like core domain and the 15 il helical domain are introduced outside of known structural elements like alpha-helices.

Hybrid A     G $\alpha$ S1-67/GPA66-299/G $\alpha$ S203-394  
This hybrid contains the entire il insert of GPA1 interposed into  
the G $\alpha$ S sequence.

20              Hybrid B     GPA1-41/G $\alpha$ S4443-67/GPA66-299/G $\alpha$ S203-394  
This hybrid contains the amino terminal 41 residues of GPA1 in  
place of the 42 amino terminal residues of G $\alpha$ S found in  
Hybrid A.

25              iii. G $\alpha$ s Hybrids.

There is evidence that the "switch region" encoded by residues 171-237 of G $\alpha$  transducin (using the numbering of (Noel et al (supra) also plays a role in G $\beta\gamma$  coupling. First, the G226A mutation in G $\alpha$ S prevents the GTP-induced conformational change that occurs with exchange of GDP for GTP upon receptor activation by ligand. 30 This residue maps to the highly conserved sequence -DVGGQ- (SEQ ID NO: 5), present in all G $\alpha$  subunits and is involved in GTP hydrolysis. In both the G $\alpha$ t and G $\alpha$  il crystal structures, this sequence motif resides in the loop that connects the  $\beta$ 3 sheet and the  $\alpha$ 2 helix in the guanine nucleotide binding core. In addition to blocking the conformational change that occurs upon GTP binding, this mutation also prevents 35 dissociation of GTP-ligated G $\alpha$ s from G $\beta\gamma$ . Second, crosslinking data reveals that a highly conserved cysteine residue in the  $\alpha$ 2 helix (C215 in G $\alpha$ o, C210 in G $\alpha$ t) can be

- 45 -

crosslinked to the carboxy terminal region of G $\beta$  subunits. Finally, genetic evidence (Whiteway et al. (1993) Mol Cell Biol. 14:3233-3239) identifies an important single residue in GPA1 (E307) in the  $\beta$ 2 sheet of the core structure that may be in direct contact with G $\gamma$ . A mutation in the GPA1 protein at this position suppresses the 5 constitutive signaling phenotype of a variety of STE4 (G $\beta$ ) dominant negative mutations that are also known to be defective in G $\alpha$ -G $\beta$  $\gamma$  association (as assessed in two-hybrid assay in yeast as well as by more conventional genetic tests).

The GPA1 switch region suppresses coupling to yeast G $\beta$  $\gamma$  (SGS), while in the context of the GPA1 amino terminus the GPA1 switch region stabilizes coupling 10 with G $\beta$  $\gamma$  (GP $\beta$  $\gamma$ -SGS). This suggests that these two regions of GPA1 collaborate to allow interactions between G $\alpha$  subunits and G $\beta$  $\gamma$  subunits. This conclusion is somewhat mitigated by the observation that the GPA41-G $\alpha$ S hybrid that does not contain the GPA1 switch region is able to complement the growth arrest phenotype of gpal strains.

The role of the surface-exposed residues of this region may be crucial for 15 effective coupling to yeast G $\beta$  $\gamma$ , and can be incorporated into hybrid molecules as follows below.

G $\alpha$ S-GPA-Switch    GQS 1-202/GPA298-350/G $\alpha$ S 253-394  
This hybrid encodes the entire switch region of GPA 1 in the context of 20 G $\alpha$ S.

G $\alpha$ S-GPA- $\alpha$ 2    GQS 1-226/GPA322-332/GQS 238-394  
This hybrid encodes the  $\alpha$ 2 helix of GPA1 in the context of G $\alpha$ S.

GPA41-G $\alpha$ S-GPA- $\alpha$ 2GPA1-41/GQS43-226/GPA322-332/GQS238-394  
This hybrid encodes the 41 residue amino terminal domain of GPA1 and 25 the  $\alpha$ 2 helix of GPA1 in the context of G $\alpha$ S.

In addition, hybrids that alter the surface exposed residues of the  $\beta$ 2 and 30  $\beta$ 3 sheets of  $\alpha$ S so that they resemble those of the GPA1 QS helix can be made. These altered  $\alpha$ 2 helical domains have the following structure. (The positions of the altered residues correspond to G $\alpha$ S.)

L203K  
K211E  
D215G  
K216S

- 46 -

### D229S

These single mutations can be engineered into a G $\alpha$ S backbone singly and in pairwise combinations. In addition, they can be introduced in the context of both the full length G $\alpha$ S and the GPA41-G $\alpha$ S hybrid described previously. All are predicted 5 to improve the coupling of G $\alpha$  subunits to yeast G $\beta$  $\gamma$  subunits by virtue of improved electrostatic and hydrophobic contacts between this region and the regions of G $\beta$  defined by Whiteway and coworkers (Whiteway et al (supra) that define site(s) that interact with GPA1).

In summary, the identification of hybrid G $\alpha$  subunits that couple to the 10 yeast pheromone pathway has led to the following general observations. First. GPA<sub>BAMH1</sub> hybrids associate with yeast G $\beta$  $\gamma$ , therefore at a minimum these hybrids contain the determinants in GPA1 necessary for coupling to the pheromone response pathway. Second, the amino terminal 41 residues of GPA1 contain sufficient 15 determinants to facilitate coupling of G $\alpha$  hybrids to yeast G $\beta$  $\gamma$  in some, but not all, instances, and that some G $\alpha$  subunits contain regions outside of the first 41 residues that are sufficiently similar to those in GPA1 to facilitate interaction with GPA1 even in the absence of the amino terminal 41 residues of GPA1. Third, there are other determinants 20 in the first 310 residues of GPA1 that are involved in coupling G $\alpha$  subunits to yeast G $\beta$  $\gamma$  subunits.

The various classes of hybrids noted above are not mutually exclusive. 25 For example, a GPA1 containing GPA1-41 could also feature the L203K mutation.

While, for the sake of simplicity, hybrids of yeast GPA1 and a mammalian G $\alpha$ s have been described here, it will be appreciated that hybrids may be made of other yeast G $\alpha$  subunits and/or other mammalian G $\alpha$  subunits, notably 25 mammalian G $\alpha$ i subunits. Moreover, while the described hybrids are constructed from two parental proteins, hybrids of three or more parental proteins are also possible.

As shown in the Examples, chimeric G $\alpha$  subunits have been especially useful in coupling receptors to G $\alpha$ i species.

#### 30 iv. Expression of G $\alpha$

Kang et al. supra reported that several classes of native mammalian G $\alpha$  subunits were able to interact functionally with yeast  $\alpha$  subunits when expression of G $\alpha$  was driven from a constitutively active, strong promoter (PGK) or from a strong 35 inducible promoter (CUP). These authors reported that rat G $\alpha$ S, G $\alpha$ i2 or G $\alpha$ o expressed at high level coupled to yeast  $\beta$  $\gamma$ . High level expression of mammalian G $\alpha$  (i.e. non-stoichiometric with respect to yeast  $\beta$  $\gamma$ ) is not preferred for uses like those described

- 47 -

in this application. Reconstruction of G protein-coupled receptor signal transduction in yeast requires the signaling component of the heterotrimeric complex (G $\beta\gamma$ ) to be present stoichiometrically with G $\alpha$  subunits. An excess of G $\alpha$  subunits (as was required for coupling of mammalian G $\alpha i2$  and G $\alpha o$  to yeast G $\beta\gamma$  in Kang et al.) would dampen  
5 the signal in systems where G $\beta\gamma$  subunits transduce the signal. An excess of G $\alpha$  subunits raises the background level of signaling in the system. Preferably, levels of G $\alpha$  and G $\beta\gamma$  subunits are balanced. For example, heterologous G $\alpha$  subunits may be expressed from a low copy (CEN ARS) vector containing the endogenous yeast GPA1 promoter and the GPA1 3' untranslated region. The minimum criterion, applied to a heterologous  
10 G $\alpha$  subunit with respect to its ability to couple functionally to the yeast pheromone pathway, is that it complement a gpal genotype when expressed from the GPA1 promoter on low copy plasmids or from an integrated, single copy gene. In the work described in this application, all heterologous G $\alpha$  subunits have been assayed in two biological systems. In the first assay heterologous G $\alpha$  subunits are tested for an ability to  
15 functionally complement the growth arrest phenotype of gpal strains. In the second assay the transcription of a fusi-HIS3 reporter gene is used to measure the extent to which the pheromone response pathway is activated, and hence the extent to which the heterologous G $\alpha$  subunit sequesters the endogenous yeast G $\beta\gamma$  complex. Mammalian G $\alpha s$ , G $\alpha i2$ , G $\alpha i3$ , G $\alpha q$ , G $\alpha 11$ , G $\alpha 16$ , G $\alpha o_a$ , G $\alpha o_b$ , and G $\alpha z$  from rat, murine or human  
20 origins were expressed from a low copy, CEN ARS vector containing the GPA1 promoter. Functional complementation of gpal strains was not observed in either assay system with any of these full-length G $\alpha$  constructs with the exception of rat and human G $\alpha s$ .

25                   D. Chimeric Yeast  $\beta\gamma$  subunits

In addition to or in place of modifying G protein G $\alpha$  subunits, yeast or heterologous G $\beta$  or G $\gamma$  subunits can be modified. The methods described above with regard to G $\alpha$  modification can be used to alter either or both of these subunits as well. For example, alignments of the yeast sequence and heterologous sequences can be made  
30 and combined with information regarding important functional domains. Such information can then be used to provide guidance in making mutations in yeast or heterologous sequences. Likewise, chimeric G $\beta$  or G $\gamma$  molecules can be constructed to enhance the coupling of heterologous GPCRs to a yeast pheromone signaling pathway.

The yeast STE4 and STE18 are related to the metazoan G protein  $\beta$  and  $\gamma$  subunits, respectively (Whiteway et al. 1989. Cell. 56:467). The  $\beta$  and  $\gamma$  subunits must be capable of interaction with one another as well as with the  $\alpha$  subunit and with the

effector. Previous work has suggested that mammalian  $\beta$  or  $\gamma$  subunits are divergent enough from their yeast homologues that they cannot functionally replace STE4 or STE 18. (Coria et al. 1996. Yeast. 12:41). Thus, in preferred embodiments, modifications are made to heterologous G $\beta$  or G $\gamma$  subunits expressed in yeast and/or chimeric subunits 5 are made to enhance heterologous receptor coupling.

The primary structure of G-protein  $\beta$  subunits is highly conserved from yeast to humans: Ste4 shares approximately 40% identity with human G $\beta$  isoforms (Leberer et al. 1992 EMBO Journal 11:4085). STE 4 and the G $\beta$ s are 420. and 340 or 10 341 amino acids long, respectively, and belong to the family of proteins with WD-40 motifs (van der Voorn and Ploegh. 1992. FEBS Lett. 307:131). These motifs can be used to divide G $\beta$  and STE4 into eight blocks (Coria et al. Yeast 1996. 12:41). Among the mammalian G $\beta$ s, some have been found to exhibit G $\gamma$  subunit selectivity (Pronin and Gautham. 1992. Proc. Natl. Acad. Sci. USA 89:6220; Schmidt et al. 1992. J. Biol. Chem. 267:13807; Kleuss et al. 1992. Nature. 358:424). An alignment of the 15 metazoan and yeast G protein  $\beta$  subunits is provided by Corai et al. (1996. Yeast. 12:41). Such an alignment can be used to provide guidance for making mutations to G protein  $\beta$  subunits as described for G $\alpha$  above. In addition, certain regions of STE4 have been found to be important and thus, may be less amenable to manipulation than other portions of the polypeptide. For example, the c-terminus of the STE4 product is 20 essential for downstream signaling (Coria et al. 1995. FEBS Letters 367:122). Mutations to two small regions in the amino terminal half of Ste4 have also been shown to inhibit signaling (Leberer et al. supra). Mutations which influence the interaction of STE4 and GPA1 have also been identified; mutations to the second copy of the WD40 repeat can be modified to reduce the interaction between STE4 and GPA1, without 25 influencing other aspects of STE4 function (Whiteway et al. 1994. Mol. Cell. Biol. 14:3223)

The G $\gamma$ s, including STE18, diverge more strongly from each other than do the G $\beta$ s. Even among the mammalian G protein  $\gamma$  subunits, there is a fair amount of divergence. The  $\gamma$  subunit may determine the functional specificity of the  $\beta\gamma$  subunit complex. Complete cDNAs for the  $\gamma$ 1 subunit from bovine retina (Hurley et al. Proc. Nat'l Acad. Sci USA. 1984. 81:6948) the  $\gamma$ 1,  $\gamma$ 3, and  $\gamma$ 7 subunits from bovine brain (Robishaw et al. J. Biol. Chem. 1989. 264:15758; Gautam et al. Science. 1989. 244:971; Gautam et al. Proc. Nat'l Acad. Sci. USA. 1990 87:7973; Cali et al. J. Biol. Chem. 1992. 267:24023), and the  $\gamma$ 5 subunit from bovine and rat liver (Gisher et al. 30 35 1992. 12:1585) have been reported.

The STE18 gene of yeast terminates with a CAAX (SEQ ID NO:83) box (where A is an aliphatic amino acid, and X is any uncharged amino acid). This sequence is involved in prenylation of Gy and is likely important in the localization of Gy to the membrane and may, thus, be less amenable to manipulation than other portions of the sequence. (Kurjan. 1992. Ann. Rev. Biochem. 61:1097). Saturation mutagenesis has also provided insight into regions of STE18 that are important in STE18 function. Mutations in STE18 which compensate for mutations in STE4 were identified at serine 65, threonine 71, and valine 80. Dominant negative alleles of the STE18 gene were also identified (Whiteway et al. 1992. Biochem. Cell. Biol. 70:1230). These truncated proteins were found to lack the carboxyl terminus of STE 18, including the CAAX box (Whiteway et al. supra).

An alignment of yeast Gy, STE18, and mammalian Gys can be made as indicated for the other G protein subunits. Such an alignment can be used in constructing mutant Gy subunits or chimeric Gy subunits. Specific examples of chimeric Gy subunits are provided in the appended examples, for instance in Example 3. In preferred embodiments, mammalian Gy2 is used in making G protein  $\gamma$  subunit chimeras.

20      E. Identification Of Mutants Which Show Enhanced Coupling To  
Heterologous GPCRS

Any of a number of methods can be used to screen for mutated or chimeric G protein subunits which show enhanced coupling to the heterologously expressed GPCRs.

For example, in order to show that the mutant Gas subunits can not only associate tightly with yeast  $\beta\gamma$ , but also dissociate from  $\beta\gamma$  upon receptor stimulation, an episomal fusl-lacZ reporter plasmid can be introduced along with the heterologous receptor. The addition of an agonist should result in an increase in  $\beta$ -galactosidase units, demonstrating the ability of the mutant Gas to interact productively with receptor, and to dissociate from  $\beta\gamma$  upon ligand addition. In addition, the generation of second messengers or mating factor responses (e.g., growth arrest or shmoo formation could be measured).

In another example, such a screen can take advantage of the fact that a gpal fusl -HIS3 colony expressing wild type Gas can grow upon replica plating to media lacking histidine and containing 1mM 3-aminotriazole (AT). The growth of this strain occurs due to the partially constitutive state of the pheromone pathway, which leads to partial derepression of the fusl-HIS3 reporter gene. AT inhibits the activity of IGP

- 50 -

dehydrazinase. Cells with low, fixed level of expression of HIS3 are sensitive to the drug, while cells with higher levels are resistant. The amount of AT can be selected to inhibit cells with a basal level of HIS3 expression (whatever that level is) but allow growth of cells with an induced level of expression. A colony containing the desired  $\beta$ y-coupling 5 Gas mutant will presumably fail to grow on this media due to the inactivation of the pheromone pathway mediated by tight  $\alpha$ y association. However, if the selection for the Gas-bearing plasmid is relaxed, in this case by the addition of limiting amounts of adenine, then the desired mutant colony will only produce growth from that fraction of 10 cells in the colony (usually 5-10%) that have lost the Gas plasmid. Such a colony will appear red due to the buildup of an intermediate of the adenine pathway. The nonmutated Gas-containing colonies will appear white on an adenine-limiting plate, since the presence of the Gas-ADE2 plasmid has no appreciable negative effect on the ability of the cells to grow in 1mM 3-AT, but confers a selective advantage under 15 adenine-limiting conditions. The visual aspect of the screen allows the facile identification of potential mutants, and eliminates potential unwanted mutations, such as Ste-, because the screen demands a plasmid-dependent phenotype.

Plasmid DNA can then be recovered from putative mutant-bearing cells, amplified and retested on SCAH1 to confirm the plasmid linkage of the mutation. Sequencing of the mutants can then be performed. In addition, the mutations can be 20 subcloned into yeast to rule out the possibility that mutations in noncoding sequences confer the His<sup>+</sup>phenotype.

## VII. Leader Sequences

It has been demonstrated that most of the mammalian extracellular, 25 secreted proteins are poorly secreted when expressed in yeast. However, in many cases their secretion levels are markedly increased when their native signal sequences are replaced by the signal sequences of yeast proteins that interact more efficiently with the ER translocation complex. Specifically, the signal sequences of yeast invertase and acid 30 phosphatase have been widely used in biotechnology to direct the secretory expression of the heterologous proteins. However, it is well established that even though many foreign proteins are targeted to the ER by the yeast signal sequences, not all of them advance further in the secretory pathway. The major problem consists in the malfolding and/or improper glycosylation of the heterologous proteins that results in their retention 35 in the ER by the quality control apparatus of the yeast cell.

In many cases, a leader sequence of a precursor of yeast mating pheromone,  $\alpha$ -factor, have been used successfully to overcome this problem [1. Brake, A.J. (1989) in Yeast Genetic Engineering (Barr, P.J., Brake, A.J., and Valenzuela, P., eds) pp. 269-280, Butterworths, London; Brake, A.J. (1990) Meth. Enzymol. 185, 408-441., and references cited therein]. This sequence, in addition to the N-terminal signal peptide of 17 residues, includes a hydrophilic pro-region which contains 72 residues and bears three sites of N-linked glycosylation. The pro-region is extensively glycosylated in the ER and Golgi and is cleaved by Kex2 endopeptidase in the late Golgi compartment. The presence of the pro-region at the N-terminus is believed to allow some heterologous proteins to pass the quality control in the ER and to reach the periplasm. It is likely that the pro-region can somehow facilitate correct protein folding. Alternatively, it may be recognized by the quality control apparatus as a properly folded structural unit thus allowing an entire fusion protein to leave the ER.

The invertase leader can also be used. This leader sequence has been demonstrated to be cleaved from nascent invertase peptide, or nascent heterologous peptide, upon entrance into the endoplasmic reticulum, the apparent molecular weight of the receptor is consistent with this interpretation. The enzyme responsible for cleavage of the pre sequence, Kex2, resides in the trans Golgi.

#### 20 A. Peptide Expression

In certain embodiments, such a leader sequence can be used to express a peptide library of the present invention. Yeast cells are bounded by a lipid bilayer called the plasma membrane. Between this plasma membrane and the cell wall is the periplasmic space. Peptides secreted by yeast cells cross the plasma membrane through a variety of mechanisms and thereby enter the periplasmic space. The secreted peptides are then free to interact with other molecules that are present in the periplasm or displayed on the outer surface of the plasma membrane. The peptides then either undergo re-uptake into the cell, diffuse through the cell wall into the medium, or become degraded within the periplasmic space.

30 The test polypeptide library may be secreted into the periplasm by any of a number of exemplary mechanisms, depending on the nature of the expression system to which they are linked. In one embodiment, the peptide may be structurally linked to a yeast signal sequence, such as that present in the  $\alpha$ -factor precursor, which directs secretion through the endoplasmic reticulum and Golgi apparatus. Since this is the same route that the receptor protein follows in its journey to the plasma membrane, opportunity exists in cells expressing both the receptor and the peptide library for a

specific peptide to interact with the receptor during transit through the secretory pathway. This has been postulated to occur in mammalian cells exhibiting autocrine activation. Such interaction could yield activation of the response pathway during transit, which would still allow identification of those cells expressing a peptide agonist.

- 5 For situations in which peptide antagonists to externally applied receptor agonist are sought, this system would still be effective, since both the peptide antagonist and receptor would be delivered to the outside of the cell in concert. Thus, those cells producing an antagonist would be selectable, since the peptide antagonist would be properly and timely situated to prevent the receptor from being stimulated by the  
10 externally applied agonist.

An alternative mechanism for delivering peptides to the periplasmic space is to use the ATP-dependent transporters of the STE6/MDR1 class. This transport pathway and the signals that direct a protein or peptide to this pathway are not as well characterized as is the endoplasmic reticulum-based secretory pathway. Nonetheless,  
15 these transporters apparently can efficiently export certain peptides directly across the plasma membrane, without the peptides having to transit the ER/Golgi pathway. It is anticipated that at least a subset of peptides can be secreted through this pathway by expressing the library in context of the  $\alpha$ -factor prosequence and terminal tetrapeptide.  
20 The possible advantage of this system is that the receptor and peptide do not come into contact until both are delivered to the external surface of the cell. Thus, this system strictly mimics the situation of an agonist or antagonist that is normally delivered from outside the cell. Use of either of the described pathways is within the scope of the invention.

The present invention does not require periplasmic secretion of peptides.  
25 or, if such secretion is provided, any particular secretion signal or transport pathway. In certain embodiments, peptides expressed with a signal sequence may bind to and activate receptors prior to their transport to the cell surface.

#### B. GPCR Expression

30 In other embodiments, a leader sequence of a yeast secreted protein can be used to direct transport of receptors, for example, G-protein coupled receptors to the plasma membrane as described in detail in the appended examples. Previous work has demonstrated the expression of foreign, secreted proteins in yeast cells using the  $\alpha$ -factor leader. However, when a heterologous membrane bound receptor, the rat M5  
35 receptor, was expressed using such a system, it was found that the heterologous GPCR did not functionally integrate into the yeast cell signaling pathway (Huang *et al.*

- 53 -

Biochem. and Biophys. Res. Comm. 1992. 182:1180). The transport of both secreted and transmembrane proteins into the endoplasmic reticulum in yeast is promoted by the same protein translocation complex, including the Sec61, Sec62 and Sec63 proteins. All the secreted proteins possess a signal sequence at their N-termini which is recognized by the translocation complex and serves as an ER targeting signal. A typical signal sequence is comprised of several positively charged residues at the N-terminus followed by a hydrophobic core and a C-terminal site of processing by signal peptidase. Some transmembrane proteins, for example, metabotropic glutamate receptors and vasoactive intestinal polypeptide receptors, also possess the N-terminal signal sequences. whereas some do not. In the latter case, a first transmembrane domain is believed to interact with the ER translocation machinery. The use of the  $\alpha$ -factor leader sequence may, therefore, be particularly desirable for functional expression of certain receptors.

In certain embodiments, it will be desirable to further modify the yeast cells of the present invention. For example, in one embodiment it will be desirable to disrupt the yeast calnexin-like gene, CNE1, to improve receptor transport from the endoplasmic reticulum to the Golgi. In other embodiments it will be desirable to disrupt of the STP22 to improve transport form the Golgi to the plasma membrane. In yet other embodiments, it will be desirable to overexpress the gene encoding Ast1, to increase transport form the Golgi to the plasma membrane. I yet other embodiments, it will be desirable to disrupt END3 and/or END4, to prevent receptor internalization.

### VIII. Test Compounds

#### A. Exogenously added compounds

A recent trend in medicinal chemistry includes the production of mixtures of compounds, referred to as libraries. While the use of libraries of peptides is well established in the art, new techniques have been developed which have allowed the production of mixtures of other compounds, such as benzodiazepines (Bunin et al. 1992. J. Am. Chem. Soc. 114:10987; DeWitt et al. 1993. Proc. Natl. Acad. Sci. USA 90:6909) peptoids (Zuckermann. 1994. J. Med. Chem. 37:2678) oligocarbamates (Cho et al. 1993. Science. 261:1303), and hydantoins (DeWitt et al. supra). Rebek et al. have described an approach for the synthesis of molecular libraries of small organic molecules with a diversity of 104-105 (Carell et al. 1994. Angew. Chem. Int. Ed. Engl. 33:2059; Carell et al. Angew. Chem. Int. Ed. Engl. 1994. 33:2061).

The compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including:

biological libraries; spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the 'one-bead one-compound' library method, and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four  
5 approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. Anticancer Drug Des. 1997. 12:145).

In one embodiment, the test compound is a peptide or peptidomimetic. In another, preferred embodiment, the compounds are small, organic non-peptidic compounds.

10 Other exemplary methods for the synthesis of molecular libraries can be found in the art, for example in: Erb et al. 1994. Proc. Natl. Acad. Sci. USA 91:11422; Horwell et al. 1996 Immunopharmacology 33:68; and in Gallop et al. 1994. J. Med. Chem. 37:1233. In addition, libraries such as those described in the commonly owned applications U.S.S.N. 08/864,241, U.S.S.N. 08/864,240 and U.S.S.N. 08/835,623 can  
15 be used to provide compounds for testing in the present invention. The contents of each of these applications is expressly incorporated herein by this reference.

Libraries of compounds may be presented in solution (e.g., Houghten (1992) Biotechniques 13:412-421), or on beads (Lam (1991) Nature 354:82-84), chips (Fodor (1993) Nature 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner  
20 USP '409), plasmids (Cull et al. (1992) Proc Natl Acad Sci USA 89:1865-1869) or on phage (Scott and Smith (1990) Science 249:386-390); (Devlin (1990) Science 249:404-406); (Cwirla et al. (1990) Proc. Natl. Acad. Sci. 87:6378-6382); (Felici (1991) J. Mol. Biol. 222:301-310); (Ladner *supra*).

25 In certain embodiments, the test compounds are exogenously added to the yeast cells expressing a recombinant receptor and compounds that modulate signal transduction via the receptor are selected. In other embodiments, the yeast cells express the compounds to be tested. For example, a culture of the subject yeast cells can be further modified to collectively express a peptide library as described in more detail in PCT Publication WO 94/23025 the contents of which is expressly incorporated herein  
30 by this reference.

Other types of peptide libraries may also be expressed, see, for example, U.S. Patents 5,270,181 and 5,292,646; and PCT publication WO94/ 02502). In still another embodiment, the combinatorial polypeptides are produced from a cDNA library.

35 Exemplary compounds which can be screened for activity include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, and natural product extract libraries. In such embodiments, both compounds which agonize

- 55 -

or antagonize the receptor- or channel-mediated signaling function can be selected and identified.

#### B. Peptide Libraries

5 In certain embodiments, yeast cells can be engineered to produce the compounds to be tested. This assay system has the advantage of increasing the effective concentration of the compound to be tested. In one embodiment, a method such as that described in WO 94/23025 can be utilized.

10 Other methods can also be used. For example, peptide libraries are systems which simultaneously display, in a form which permits interaction with a target, a highly diverse and numerous collection of peptides. These peptides may be presented in solution (Houghten (1992) Biotechniques 13:412-421), or on beads (Lam (1991) Nature 354:82-84), chips (Fodor (1993) Nature 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull et al. (1992) Proc Natl Acad Sci USA 89:1865-1869) or on phage (Scott and Smith (1990) Science 249:386-390); (Devlin (1990) Science 249:404-406); (Cwirla et al. (1990) Proc. Natl. Acad. Sci. 87:6378-6382); (Felici (1991) J. Mol. Biol. 222:301-310); (Ladner supra.). Many of these systems are limited in terms of the maximum length of the peptide or the composition of the peptide (e.g., Cys excluded). Steric factors, such as the proximity of 20 a support, may interfere with binding. Usually, the screening is for binding in vitro to an artificially presented target, not for activation or inhibition of a cellular signal transduction pathway in a living cell. While a cell surface receptor may be used as a target, the screening will not reveal whether the binding of the peptide caused an allosteric change in the conformation of the receptor.

25 The Ladner et al. patent, USSN 5,096,815, describes a method of identifying novel proteins or polypeptides with a desired DNA binding activity. Semi-random ("variegated") DNA encoding a large number of different potential binding proteins is introduced, in expressible form, into suitable yeast cells. The target DNA sequence is incorporated into a genetically engineered operon such that the 30 binding of the protein or polypeptide will prevent expression of a gene product that is deleterious to the gene under selective conditions. Cells which survive the selective conditions are thus cells which express a protein which binds the target DNA. While it is taught that yeast cells may be used for testing, bacterial cells are preferred. The interactions between the protein and the target DNA occur only in the cell (and then only 35 in the nucleus), not in the periplasm or cytoplasm, and the target is a nucleic acid, and not a receptor protein. Substitution of random peptide sequences for functional domains

in cellular proteins permits some determination of the specific sequence requirements for the accomplishment of function. Though the details of the recognition phenomena which operate in the localization of proteins within cells remain largely unknown, the constraints on sequence variation of mitochondrial targeting sequences and protein secretion signal sequences have been elucidated using random peptides (Lemire et al., J. Biol. Chem. (1989) 264, 20206 and Kaiser et al. (1987) Science 235:312, respectively).

In certain embodiments of the instant invention, the compounds tested are in the form of peptides from a peptide library. The peptide library of the present invention takes the form of a cell culture, in which essentially each cell expresses one, and usually only one, peptide of the library. While the diversity of the library is maximized if each cell produces a peptide of a different sequence, it is usually prudent to construct the library so there is some redundancy. Depending on size, the combinatorial peptides of the library can be expressed as is, or can be incorporated into larger fusion proteins. The fusion protein can provide, for example, stability against degradation or denaturation, as well as a secretion signal if secreted. In an exemplary embodiment of a library for intracellular expression, e.g., for use in conjunction with intracellular target receptors, the polypeptide library is expressed as thioredoxin fusion proteins (see, for example, U.S. Patents 5,270,181 and 5,292,646; and PCT publication WO94/ 02502). The combinatorial peptide can be attached one the terminus of the thioredoxin protein, or, for short peptide libraries, inserted into the so-called active loop.

In one embodiment, the peptide library is derived to express a combinatorial library of polypeptides which are not based on any known sequence, nor derived from cDNA. That is, the sequences of the library are largely random. In preferred embodiments, the combinatorial polypeptides are in the range of 3-100 amino acids in length, more preferably at least 5-50, and even more preferably at least 10, 13, 15, 20 or 25 amino acid residues in length. Preferably, the polypeptides of the library are of uniform length. It will be understood that the length of the combinatorial peptide does not reflect any extraneous sequences which may be present in order to facilitate expression, e.g., such as signal sequences or invariant portions of a fusion protein.

In another embodiment, the peptide library is a combinatorial library of polypeptides which are based at least in part on a known polypeptide sequence or a portion thereof (not a cDNA library). That is, the sequences of the library is semi-random, being derived by combinatorial mutagenesis of a known sequence. See, for example, Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) J. Biol. Chem. 267:16007-16010; Griffiths et al. (1993) EMBO J 12:725-734; Clackson et al. (1991) Nature 352:624-628; and Barbas et

al. (1992) PNAS 89:4457-4461. Accordingly, polypeptide(s) which are known ligands for a target receptor can be mutagenized by standard techniques to derive a variegated library of polypeptide sequences which can further be screened for agonists and/or antagonists. For example, the surrogate ligand identified for FPRL-1, e.g., the Ser-Leu-  
5 Leu-Trp-Leu-Thr-Cys-Arg-Pro-Trp-Glu-Ala-Met peptide, can be mutagenized to generate a library of peptides with some relationship to the original tridecapeptide. This library can be expressed in a reagent cell of the present invention, and other receptor activators can be isolated from the library. This may permit the identification of even more potent FPRL-1 surrogate ligands.

10 In still another embodiment, the combinatorial polypeptides are produced from a cDNA library.

In a preferred embodiment of the present invention, the yeast cells collectively produce a "peptide library", preferably including at least  $10^3$  to  $10^7$  different peptides, so that diverse peptides may be simultaneously assayed for the ability to  
15 interact with the exogenous receptor. In an especially preferred embodiment, at least some peptides of the peptide library are secreted into the periplasm, where they may interact with the "extracellular" binding site(s) of an exogenous receptor. They thus mimic more closely the clinical interaction of drugs with cellular receptors. This embodiment optionally may be further improved (in assays not requiring pheromone  
20 secretion) by preventing pheromone secretion, and thereby avoiding competition between the peptide and the pheromone for signal peptidase and other components of the secretion system.

In certain embodiments of the present invention, the peptides of the library are encoded by a mixture of DNA molecules of different sequence. Each  
25 peptide-encoding DNA molecule is ligated with a vector DNA molecule and the resulting recombinant DNA molecule is introduced into a yeast cell. Since it is a matter of chance which peptide encoding DNA molecule is introduced into a particular cell, it is not predictable which peptide that cell will produce. However, based on a knowledge of the manner in which the mixture was prepared, one may make certain statistical  
30 predictions about the mixture of peptides in the peptide library.

The peptides of the library can be composed of constant and variable residues. If the nth residue is the same for all peptides of the library, it is said to be constant. If the nth residue varies, depending on the peptide in question, the residue is a variable one. The peptides of the library will have at least one, and usually more than  
35 one, variable residue. A variable residue may vary among any of two to all twenty of the genetically encoded amino acids; the variable residues of the peptide may vary in the

same or different manner. Moreover, the frequency of occurrence of the allowed amino acids at a particular residue position may be the same or different. The peptide may also have one or more constant residues.

There are two principal ways in which to prepare the required DNA mixture. In one method, the DNAs are synthesized a base at a time. When variation is desired, at a base position dictated by the Genetic Code, a suitable mixture of nucleotides is reacted with the nascent DNA, rather than the pure nucleotide reagent of conventional polynucleotide synthesis.

The second method provides more exact control over the amino acid variation. First, trinucleotide reagents are prepared, each trinucleotide being a codon of one (and only one) of the amino acids to be featured in the peptide library. When a particular variable residue is to be synthesized, a mixture is made of the appropriate trinucleotides and reacted with the nascent DNA. Once the necessary "degenerate" DNA is complete, it must be joined with the DNA sequences necessary to assure the expression of the peptide, as discussed in more detail below, and the complete DNA construct must be introduced into the yeast cell.

In embodiments in which the test compounds it may be desirable to express such peptides in the context of a leader sequence. Yeast cells are bounded by a lipid bilayer called the plasma membrane. Between this plasma membrane and the cell wall is the periplasmic space. Peptides secreted by yeast cells cross the plasma membrane through a variety of mechanisms and thereby enter the periplasmic space. The secreted peptides are then free to interact with other molecules that are present in the periplasm or displayed on the outer surface of the plasma membrane. The peptides then either undergo re-uptake into the cell, diffuse through the cell wall into the medium, or become degraded within the periplasmic space.

The test polypeptide library may be secreted into the periplasm by any of a number of exemplary mechanisms, depending on the nature of the expression system to which they are linked. In one embodiment, the peptide may be structurally linked to a yeast signal sequence, such as that present in the  $\alpha$ -factor precursor, which directs secretion through the endoplasmic reticulum and Golgi apparatus. Since this is the same route that the receptor protein follows in its journey to the plasma membrane, opportunity exists in cells expressing both the receptor and the peptide library for a specific peptide to interact with the receptor during transit through the secretory pathway. This has been postulated to occur in mammalian cells exhibiting autocrine activation. Such interaction could yield activation of the response pathway during transit, which would still allow identification of those cells expressing a peptide agonist.

For situations in which peptide antagonists to externally applied receptor agonist are sought, this system would still be effective, since both the peptide antagonist and receptor would be delivered to the outside of the cell in concert. Thus, those cells producing an antagonist would be selectable, since the peptide antagonist would be properly and timely situated to prevent the receptor from being stimulated by the externally applied agonist.

- An alternative mechanism for delivering peptides to the periplasmic space is to use the ATP-dependent transporters of the STE6/MDR1 class. This transport pathway and the signals that direct a protein or peptide to this pathway are not as well characterized as is the endoplasmic reticulum-based secretory pathway. Nonetheless, these transporters apparently can efficiently export certain peptides directly across the plasma membrane, without the peptides having to transit the ER/Golgi pathway. It is anticipated that at least a subset of peptides can be secreted through this pathway by expressing the library in context of the a-factor prosequence and terminal tetrapeptide.
- The possible advantage of this system is that the receptor and peptide do not come into contact until both are delivered to the external surface of the cell. Thus, this system strictly mimics the situation of an agonist or antagonist that is normally delivered from outside the cell. Use of either of the described pathways is within the scope of the invention. The present invention does not require periplasmic secretion, or, if such secretion is provided, any particular secretion signal or transport pathway.

#### IX. Screening and Selection: Assays of Second Messenger Generation

When screening for bioactivity of compounds, intracellular second messenger generation can be measured directly. A variety of intracellular effectors have been identified as being G-protein-regulated, including adenylyl cyclase, cyclic GMP, phosphodiesterases, phosphoinositidase C, and phospholipase A<sub>2</sub>. In addition, G proteins interact with a range of ion channels and are able to inhibit certain voltage-sensitive Ca<sup>++</sup> transients, as well as stimulating cardiac K<sup>+</sup> channels.

In one embodiment, the GTPase enzymatic activity by G proteins can be measured in plasma membrane preparations by determining the breakdown of  $\gamma^{32}\text{P}$  GTP using techniques that are known in the art (For example, see *Signal Transduction: A Practical Approach*. G. Milligan, Ed. Oxford University Press, Oxford England). When receptors that modulate cAMP are tested, it will be possible to use standard techniques for cAMP detection, such as competitive assays which quantitate [<sup>3</sup>H]cAMP in the presence of unlabelled cAMP.

Certain receptors stimulate the activity of phospholipase C which stimulates the breakdown of phosphatidylinositol 4,5 bisphosphate to 1,4,5-IP<sub>3</sub> (which mobilizes intracellular Ca<sup>++</sup>) and diacylglycerol (DAG) (which activates protein kinase C). Inositol lipids can be extracted and analyzed using standard lipid extraction techniques. DAG can also be measured using thin-layer chromatography. Water soluble derivatives of all three inositol lipids (IP<sub>1</sub>, IP<sub>2</sub>, IP<sub>3</sub>) can also be quantitated using radiolabelling techniques or HPLC.

The mobilization of intracellular calcium or the influx of calcium from outside the cell can be measured using standard techniques. The choice of the appropriate calcium indicator, fluorescent, bioluminescent, metallochromic, or Ca<sup>++</sup>-sensitive microelectrodes depends on the cell type and the magnitude and time constant of the event under study (Borle (1990) *Environ Health Perspect* 84:45-56). As an exemplary method of Ca<sup>++</sup> detection, cells could be loaded with the Ca<sup>++</sup>sensitive fluorescent dye fura-2 or indo-1, using standard methods, and any change in Ca<sup>++</sup> measured using a fluorometer.

The other product of PIP<sub>2</sub> breakdown, DAG can also be produced from phosphatidyl choline. The breakdown of this phospholipid in response to receptor-mediated signaling can also be measured using a variety of radiolabelling techniques.

The activation of phospholipase A<sub>2</sub> can easily be quantitated using known techniques, including, for example, the generation of arachidonate in the cell.

In the case of certain receptors, it may be desirable to screen for changes in cellular phosphorylation. Such assay formats may be useful when the receptor of interest is a receptor tyrosine kinase. For example, yeast transformed with the FGF receptor and a ligand which binds the FGF receptor could be screened using colony immunoblotting (Lyons and Nelson (1984) *Proc. Natl. Acad. Sci. USA* 81:7426-7430) using anti-phosphotyrosine. In addition, tests for phosphorylation could be useful when a receptor which may not itself be a tyrosine kinase, activates protein kinases that function downstream in the signal transduction pathway. Likewise, it is noted that protein phosphorylation also plays a critical role in cascades that serve to amplify signals generated at the receptor. Multi-kinase cascades allow not only signal amplification but also signal divergence to multiple effectors that are often cell-type specific, allowing a growth factor to stimulate mitosis of one cell and differentiation of another.

One such cascade is the MAP kinase pathway that appears to mediate both mitogenic, differentiation and stress responses in different cell types. Stimulation of growth factor receptors results in Ras activation followed by the sequential activation of c-Raf, MEK, and p44 and p42 MAP kinases (ERK1 and ERK2). Activated MAP kinase then phosphorylates many key regulatory proteins, including p90RSK and Elk-1 that are phosphorylated when MAP kinase translocates to the nucleus. Homologous pathways exist in mammalian and yeast cells. For instance, an essential part of the *S. cerevisiae* pheromone signaling pathway is comprised of a protein kinase cascade composed of the products of the STE11, STE7, and FUS3/KSS1 genes (the latter pair are distinct and functionally redundant). Accordingly, phosphorylation and/or activation of members of this kinase cascade can be detected and used to quantitate receptor engagement. Phosphotyrosine specific antibodies are available to measure increases in tyrosine phosphorylation and phospho-specific antibodies are commercially available (New England Biolabs, Beverly, MA).

Modified methods for detecting receptor-mediated signal transduction exist and one of skill in the art will recognize suitable methods that may be used to substitute for the example methods listed.

In one embodiment, the indicator gene can be used for detection. In one embodiment an indicator gene is an unmodified endogenous gene. For example, the instant method can rely on detecting the transcriptional level of such pheromone system pathway responsive endogenous genes as the *Bar1* or *Fus1*, *Fus2*, mating factor, *Ste3* *Ste13*, *Kex1*, *Ste2*, *Ste6*, *Ste7*, *sSst2*, or *Chs1*. (Appeltauer and Zchstetter. 1989. Eur. J. Biochem. 181:243)

In other embodiments, the sensitivity of an endogenous indicator gene can be enhanced by manipulating the promoter sequence at the natural locus for the indicator gene. Such manipulation may range from point mutations to the endogenous regulatory elements to gross replacement of all or substantial portions of the regulatory elements. The previous discussion of mutations with regard to G proteins and G protein coupled receptors is reiterated here.

For example, in the case of the *Bar1* gene, the promoter of the gene can be modified to enhance the transcription of *Bar1* upon activation of the yeast pheromone system pathway. *Bar1* gene transcription is inactivated upon exposure of yeast cells to mating factor. The sequence of the *Bar1* gene is known in the art (see e.g., U.S. patent 4,613,572). Moreover, the sequences required for  $\alpha$ -factor-enhanced expression of the

Bar1, and other pheromone responsive genes have been identified. (Appeltauer and Achstetter 1989. Eur. J. Biochem. 181:243; Hagen et al. 1991. Mol. Cell. Biol. 11:2952). In an exemplary embodiment, the yeast *Bar1* promoter can be engineered by mutagenesis to be more responsive, e.g., to more strongly promoter gene transcription, 5 upon stimulation of the yeast pheromone pathway. Standard techniques for mutagenizing the promoter can be used. In such embodiments, it is desirable that the conserved oligonucleotide motif described by Appeltauer et al. be conserved.

In yet other embodiments, rather than measuring second messenger production or alterations in transcription, the activity of endogenous yeast proteins can 10 be assayed. For example, in one embodiment, the signal transduction pathway of the receptor upregulates expression or otherwise activates an enzyme which is capable of modifying a substrate which can be added to the cell. The signal can be detected by using a detectable substrate, in which case loss of the substrate signal is monitored, or alternatively, by using a substrate which produces a detectable product. In certain 15 embodiments, the substrate is naturally occurring. Alternatively, the substrate can be non-naturally occurring. In preferred embodiments, BAR1 activity can be measured.

In other embodiments, the modulation of a receptor by a test compound can result in a change in the transcription of a gene, which is not normally pheromone responsive. In preferred embodiments, the gene is easily detectable. For example, in a 20 preferred embodiment, the subject assay can be used to measure Pho5, a secreted acid phosphatase. Acid phosphatase activity can be measured using standard techniques.

In other embodiments, reporter gene constructs can be used. Reporter gene constructs are prepared by operatively linking a reporter gene with at least one transcriptional regulatory element. If only one transcriptional regulatory element is 25 included it must be a regulatable promoter. At least one of the selected transcriptional regulatory elements must be indirectly or directly regulated by the activity of the selected cell-surface receptor whereby activity of the receptor can be monitored via transcription of the reporter genes.

Many reporter genes and transcriptional regulatory elements are known to 30 those of skill in the art and others may be identified or synthesized by methods known to those of skill in the art. Reporter genes include any gene that expresses a detectable gene product, which may be RNA or protein. Preferred reporter genes are those that are readily detectable. The reporter gene may also be included in the construct in the form of

a fusion gene with a gene that includes desired transcriptional regulatory sequences or exhibits other desirable properties.

Examples of reporter genes include, but are not limited to CAT (chloramphenicol acetyl transferase) (Alton and Vapnek (1979), Nature 282: 864-869) 5 luciferase, and other enzyme detection systems, such as beta-galactosidase; firefly luciferase (deWet et al. (1987), Mol. Cell. Biol. 7:725-737); bacterial luciferase (Engebrecht and Silverman (1984), PNAS 1: 4154-4158; Baldwin et al. (1984), Biochemistry 23: 3663-3667); alkaline phosphatase (Toh et al. (1989) Eur. J. Biochem. 182: 231-238, Hall et al. (1983) J. Mol. Appl. Gen. 2: 101), human placental secreted 10 alkaline phosphatase (Cullen and Malim (1992) Methods in Enzymol. 216:362-368) and green fluorescent protein (U.S. patent 5,491,084; WO96/23898).

Preferred reporter gene constructs include Fus1p-LacZ and Fus1p-GFP, in which the Fus1 promoter, which is pheromone responsive, is operatively linked to sequences encoding either β-galactosidase (LacZ) or green fluorescent protein (GFP). 15 The GFP sequences can be wildtype or can be mutant sequences that show altered or increased fluorescence (also referred to in the art as "enhanced GFP"). The construction and use of such Fus1p reporter gene constructs is described in further detail in the Examples. The Fus1p-GFP construct is particularly preferred as it allows for rapid, simple and sensitive detection of reporter gene expression and it can be used to control 20 for visual uniformity or variation in the signal within a population of cells.

Transcriptional control elements include, but are not limited to, promoters, enhancers, and repressor and activator binding sites. Suitable transcriptional regulatory elements may be derived from the transcriptional regulatory regions of genes whose expression is rapidly induced, generally within minutes, of contact between the 25 cell surface protein and the effector protein that modulates the activity of the cell surface protein. Examples of such genes include, but are not limited to, the immediate early genes (see, Sheng et al. (1990) Neuron 4: 477-485), such as c-fos, Immediate early genes are genes that are rapidly induced upon binding of a ligand to a cell surface protein. The transcriptional control elements that are preferred for use in the gene constructs include 30 transcriptional control elements from immediate early genes, elements derived from other genes that exhibit some or all of the characteristics of the immediate early genes, or synthetic elements that are constructed such that genes in operative linkage therewith exhibit such characteristics. The characteristics of preferred genes from which the transcriptional control elements are derived include, but are not limited to, low or 35 undetectable expression in quiescent cells, rapid induction at the transcriptional level

within minutes of extracellular simulation, induction that is transient and independent of new protein synthesis, subsequent shut-off of transcription requires new protein synthesis, and mRNAs transcribed from these genes have a short half-life. It is not necessary for all of these properties to be present.

5 Other promoters and transcriptional control elements, in addition to those described above, include the vasoactive intestinal peptide (VIP) gene promoter (cAMP responsive; Fink et al. (1988), Proc. Natl. Acad. Sci. 85:6662-6666); the somatostatin gene promoter (cAMP responsive; Montminy et al. (1986), Proc. Natl. Acad. Sci. 83:6682-6686); the proenkephalin promoter (responsive to cAMP, nicotinic agonists, 10 and phorbol esters; Comb et al. (1986), Nature 323:353-356); the phosphoenolpyruvate carboxy-kinase gene promoter (cAMP responsive; Short et al. (1986), J. Biol. Chem. 261:9721-9726); the NGFI-A gene promoter (responsive to NGF, cAMP, and serum; Changelian et al. (1989). Proc. Natl. Acad. Sci. 86:377-381); and others that may be known to or prepared by those of skill in the art.

15 In certain assays it may be desirable to use changes in growth in the screening procedure. For example, one of the consequences of activation of the pheromone signal pathway in wild-type yeast is growth arrest. If one is testing for an antagonist of a G protein-coupled receptor, this normal response of growth arrest can be used to select cells in which the pheromone response pathway is inhibited. That is, cells 20 exposed to both a known agonist and a peptide of unknown activity will be growth arrested if the peptide is neutral or an agonist, but will grow normally if the peptide is an antagonist. Thus, the growth arrest response can be used to advantage to discover peptides that function as antagonists.

25 However, when searching for compounds which can function as agonists of G protein-coupled receptors, or other pheromone system proteins, the growth arrest consequent to activation of the pheromone response pathway is an undesirable effect since cells that bind agonists stop growing while surrounding cells that fail to bind agonists will continue to grow. The cells of interest, then, will be overgrown or their detection obscured by the background cells, confounding identification of the cells of 30 interest. To overcome this problem the present invention teaches engineering the cell such that: 1) growth arrest does not occur as a result of exogenous signal pathway activation (e.g., by inactivating the FAR1 gene); and/or 2) a selective growth advantage is conferred by activating the pathway (e.g., by transforming an auxotrophic mutant with a HIS3 gene under the control of a pheromone-responsive promoter, and applying 35 selective conditions).

Alternatively, the promoter may be one which is repressed by the receptor pathway, thereby preventing expression of a product which is deleterious to the cell. With a receptor repressed promoter, one screens for agonists by linking the promoter to a deleterious gene, and for antagonists, by linking it to a beneficial gene.

- 5 Repression may be achieved by operably linking a receptor-induced promoter to a gene encoding mRNA which is antisense to at least a portion of the mRNA encoded by the marker gene (whether in the coding or flanking regions), so as to inhibit translation of that mRNA. Repression may also be obtained by linking a receptor-induced promoter to a gene encoding a DNA binding repressor protein, and incorporating a suitable operator site into the promoter or other suitable region of the marker gene.
- 10

In the case of yeast, exemplary positively selectable (beneficial) genes include the following: *URA3*, *LYS2*, *HIS3*, *LEU2*, *TRP1*; *ADE1.2.3.4.5.7.8*; *ARG1*, 3, 4, 5, 6, 8; *HIS1*, 4, 5; *ILV1*, 2, 5; *THR1*, 4; *TRP2*, 3, 4, 5; *LEU1*, 4;

15 *MET2.3.4.8.9.14.16.19*; *URA1.2.4.5.10*; *HOM3.6*; *ASP3*; *CHO1*; *ARO 2.7*; *CYS3*; *OLE1*; *JN01.2.4*; *PR01.3*

- Countless other genes are potential selective markers. The above are involved in well-characterized biosynthetic pathways. The imidazoleglycerol phosphate dehydratase (IGP dehydratase) gene (*HIS3*) is preferred because it is both quite sensitive and can be selected over a broad range of expression levels. In the simplest case, the cell is auxotrophic for histidine (requires histidine for growth) in the absence of activation. Activation leads to synthesis of the enzyme and the cell becomes prototrophic for histidine (does not require histidine). Thus the selection is for growth in the absence of histidine. Since only a few molecules per cell of IGP dehydratase are required for histidine prototrophy, the assay is very sensitive.
- 20

- In another version of the assay, cells can be selected for resistance to 25 aminotriazole (AT), a drug that inhibits the activity of IGP dehydratase. Cells with low, fixed level of expression of *HIS3* are sensitive to the drug, while cells with higher levels are resistant. The amount of AT can be selected to inhibit cells with a basal level of *HIS3* expression (whatever that level is) but allow growth of cells with an induced level of expression. In this case selection is for growth in the absence of histidine and in the presence of a suitable level of AT.
- 30

- In appropriate assays, so-called counterselectable or negatively selectable genes may be used. Suitable genes include: *URA3* (orotidine-5'-phosphate decarboxylase; inhibits growth on 5-fluoroorotic acid), *LYS2* (2-amino adipate reductase; inhibits growth on  $\alpha$ -amino adipate as sole nitrogen source), *CYH2* (encodes 35 ribosomal protein L29; cycloheximide-sensitive allele is dominant to resistant allele),

- 66 -

CAN1 (encodes arginine permease; null allele confers resistance to the arginine analog canavanine), and other recessive drug-resistant markers.

In one example, the reporter gene effects yeast cell growth. The natural response to signal transduction via the yeast pheromone system response pathway is for 5 cells to undergo growth arrest. This is a preferred way to select for antagonists of a ligand/receptor pair that stimulates a the pathway. An antagonist would inhibit the activation of the pathway; hence, the cell would be able to grow. Thus, the FAR1 gene may be considered an endogenous counterselectable marker. The FAR1 gene is preferably inactivated when screening for agonist activity.

- 10        The reporter gene may also be a screenable gene. The screened characteristic may be a change in cell morphology, metabolism or other screenable features. Suitable markers include beta-galactosidase (Xgal, C<sub>12</sub>FDG, Salmon-gal, Magenta-Gal (latter two from Biosynth Ag)), alkaline phosphatase, horseradish peroxidase, exo-glucanase (product of yeast exbl gene; nonessential, secreted); 15 luciferase; bacterial green fluorescent protein; (human placental) secreted alkaline phosphatase (SEAP); and chloramphenicol transferase (CAT). Some of the above can be engineered so that they are secreted (although not  $\beta$ -galactosidase). A preferred screenable marker gene is beta-galactosidase; yeast cells expressing the enzyme convert the colorless substrate Xgal into a blue pigment. Again, the promoter may be 20 receptor-induced or receptor-inhibited.

### XI. Carboxy-terminal chimeric G protein subunits

Proper signal transduction requires that each GPCR interact only with specific

- 5 subclasses of several structurally similar G proteins (see e.g., Conklin *et al.*, (1993) *Cell* 73, 631-641; Savarese *et al.*, (1992) *Biochem* 283, 1-19; Headin *et al.*, (1993) *Cell signalling* 5, 505-518). Correct sorting of intracellular signals depends on the specificity  
10 of GPCR/G protein interactions. To improve the coupling specificity of heterologous G protein-coupled receptors expressed in yeast, carboxy terminal chimeric G proteins were produced and tested. Chimeric G $\alpha$  proteins were designed to contain the majority of the Gpa1 sequence and a portion of a heterologous (e.g., mammalian) G-protein subunit replacing a C-terminal end of Gpa1.

Construction of chimeric G $\alpha$  proteins involves replacing C-terminal amino acids of Gpa1 with C-terminal amino acids of a heterologous (e.g., mammalian) G $\alpha$  protein

- 15 subunit. Standard molecular biology techniques can be used to obtain nucleic acid molecules encoding Gpa1 and the other heterologous G $\alpha$  subunits and construct the chimeric G $\alpha$  proteins, for example by standard PCR amplification and mutagenesis, followed by cloning into an appropriate yeast expression vector. The entire coding region of GPA1, aligned with G $\alpha$ s, G $\alpha$ i, G $\alpha$ o, G $\alpha$ q and G $\alpha$ z, is described in Dietzel and Kurjan (1987, *Cell* 50:573) and Lambright, et al. (1994, *Nature* 369:621-628). Additional G $\alpha$  subunit sequence information is provided by Mattera, et al. (1986, *FEBS Lett* 206:36-41), Bray, et al. (1986, *Proc. Natl. Acad. Sci USA* 83:8893-8897) and Bray, et al. (1987, *Proc Natl. Acad Sci USA* 84:5115-5119). An alignment of GPA1 and four other G $\alpha$  proteins is provided by Stone and Reed (1990. *Mol. Cell Biol.* 10:4439).  
20 General methods and procedures for making chimeric G proteins are described, for example in Conklin *et al.*, (1995) *Nature* 363, 274-276 and Liu *et al.*, (1995) *Proc. Natl. Acad. Sci.* 92, 111642-1116460.

The Gpa1 protein comprises amino acid positions 1-472 (Dietzel and Kurjan

- 30 (1987) *supra*). C-terminal amino acids of the Gpa1 protein can be deleted and replaced with C-terminal amino acid residues of a heterologous (e.g., mammalian) G $\alpha$  subunit, for example, the last 4 C-terminal amino acid residues, i.e., positions 469-472 of Gpa1, can be removed and replaced with C-terminal amino acid residues from a heterologous G $\alpha$  subunit. In another embodiment, the last 5 C-terminal amino acid residues, i.e. positions 468-472 of Gpa1, can be removed and replaced with C-terminal amino acid

residues from a heterologous G  $\alpha$  subunit. In another embodiment, the last 6 C-terminal amino acid residues, i.e. positions 467-472 of G $\alpha$ 1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In another embodiment, the last 28 amino acid residues, i.e. positions 445-472 of G $\alpha$ 1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In another embodiment, the last 30 amino acid residues, i.e. positions 443-472 of G $\alpha$ 1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In yet another embodiment, the last 50 amino acid residues, i.e. positions 423-472 of G $\alpha$ 1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In various other embodiments, any number of C-terminal G $\alpha$ 1 amino acid residues from 4 to 50 amino acids can be replaced, e.g., 5-10, 10-20, 20-30, 30-40, 40-50, more preferably, 5 or 6.

The C-terminal amino acids of the G $\alpha$ 1 are replaced with the C-terminal amino acids from a heterologous G  $\alpha$  subunit. For example, the last 5 C-terminal amino acids of G $\alpha$ 1, i.e., positions 468-472, are removed and replaced with the last 5 C-terminal amino acids of a heterologous G  $\alpha$  subunit. The C-terminal amino acids of the G $\alpha$ 1 amino acids can be replaced with C-terminal amino acids of essentially any heterologous G protein  $\alpha$  subunit, but most preferably are replaced with mammalian sequences, most preferably mouse, rat or human sequences. The heterologous G  $\alpha$  subunits can be selected from any member of the G  $\alpha$  subunit family, for example G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3, G $\alpha$ q, G $\alpha$ o<sub>a</sub>, G $\alpha$ o<sub>b</sub> and G $\alpha$ 16. Preferably, the G  $\alpha$  subunit is selected from G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3, G $\alpha$ q, and G $\alpha$ 16.

In one embodiment, from the last 4 to the last 50 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In another embodiment, from the last 4 to the last 40 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In yet another embodiment, from the last 4 to the last 28 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In a preferred embodiment, from the last 4 to the last 6 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In a more preferred embodiment, the last 5 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In various other embodiments, any number of C-terminal amino acid residues of the heterologous G  $\alpha$  subunit from 4 to 50 amino acids can replace G $\alpha$ 1 sequences, e.g., 4-50, 4-30, 4-40, 4-28, 4-6, more preferably, 5 or 6.

35

## SUBSTITUTE SHEET (RULE 26)

- 69 -

The G<sub>p</sub>α1 sequence preferably is replaced with a wild type sequence of a heterologous G α subunit, for example by replacing the last 5 C-terminal amino acids of the G<sub>p</sub>α1 with the corresponding last 5 C-terminal amino acids of a wildtype mammalian G α subunit. However, the G<sub>p</sub>α1 sequence may also be replaced with sequences of a mutant G α subunit containing, for example, one or more conservative substitutions in the last 5 C-terminal amino acid sequence of the G α subunit. Mutant G α subunit sequences are selected based on their ability to retain, or improve, receptor specificity. Accordingly, the C-terminal amino acids of the heterologous G<sub>α</sub> subunit are intended to include wildtype sequences and conservative substitutions that maintain receptor specificity.

Coupling specificity of the chimeric G<sub>α</sub> proteins can be tested by expressing the chimeric G<sub>α</sub> proteins in yeast cells that also express a heterologous G protein-coupled receptor. Coupling of the chimeric G<sub>α</sub> protein to a heterologous GPCR can be determined by, for example, Fus1-LacZ screening, as described in the Example 11.

Chimeric G<sub>α</sub> proteins in which the last 5 or the last 6 C-terminal amino acids of G<sub>p</sub>α1 are replaced by the last 5 or the last 6, respectively, C-terminal amino acids of a heterologous G α subunit demonstrate an improved coupling with several heterologous G protein-coupled receptors (see Tables 3 and 4 in Example 11). These chimeric G proteins are also able to couple and stimulate orphan receptors (see Table 5 in Example 11).

### XII. "Sandwich" Chimera G Proteins

Further improvements to the coupling specificity of heterologous receptors expressed in yeast, can be made by generating " sandwich chimera", in which G<sub>p</sub>α1 sequences are flanked at both the amino- and carboxy-terminal ends by heterologous G<sub>α</sub> sequences. The term "sandwich chimera" as used herein refers to a chimeric G<sub>p</sub>α1 subunit with at least one C-terminal amino acid of a first heterologous G protein α subunit and at least one N-terminal amino acid of a second heterologous G protein α subunit. The first and the second heterologous G α subunits can be the same or different, e.g., G<sub>αs</sub>-G<sub>p</sub>α1-G<sub>αs</sub> or G<sub>αs</sub>-G<sub>p</sub>α1-G<sub>αq</sub>. Sandwich chimeric G proteins were designed to contain a majority of the G<sub>p</sub>α1 sequence and a portion of a G α subunit at the C-terminal and N-terminal ends of the G<sub>p</sub>α1 protein.

Construction of sandwich chimeric G proteins typically involves replacing C-terminal amino acids of G<sub>p</sub>α1 with C-terminal amino acids of a heterologous G α

subunit and either replacing N-terminal amino acids of G<sub>p</sub>a1 with N-terminal amino acids of a heterologous G  $\alpha$  subunit or adding N-terminal amino acids of a heterologous G  $\alpha$  subunit to the amino-terminal end of G<sub>p</sub>a1. Standard molecular biology techniques can be used to obtain nucleic acid molecules encoding G<sub>p</sub>a1 and heterologous G  $\alpha$  subunit proteins. Constructs of sandwich chimeric G proteins, can be made for example, by standard PCR amplification and mutagenesis, followed by cloning into an appropriate yeast expression vector.

C-terminal amino acids of the G<sub>p</sub>a1 protein can be deleted and replaced with C-terminal amino acid residues of a heterologous (e.g., mammalian) G  $\alpha$  subunit, for example, the last 4 C-terminal amino acid residues, i.e., positions 469-472 of G<sub>p</sub>a1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In another embodiment, the last 5 C-terminal amino acid residues, i.e. positions 468-472 of G<sub>p</sub>a1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In another embodiment, the last 6 C-terminal amino acid residues, i.e. positions 467-472 of G<sub>p</sub>a1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In another embodiment, the last 28 amino acid residues, i.e., positions 445-472 of G<sub>p</sub>a1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In another embodiment, the last 30 amino acid residues, i.e., positions 443-472 of G<sub>p</sub>a1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In yet another embodiment, the last 50 amino acid residues, i.e. positions 423-472 of G<sub>p</sub>a1, can be removed and replaced with C-terminal amino acid residues from a heterologous G  $\alpha$  subunit. In various other embodiments, any number of C-terminal G<sub>p</sub>a1 amino acid residues from 4 to 50 amino acids can be replaced, e.g., 4-50, 4-40, 4-30, 4-28, 4-6, more preferably, 5 or 6.

The C-terminal amino acids of the G<sub>p</sub>a1 are replaced with the C-terminal amino acids from a heterologous G  $\alpha$  subunit. For example, the last 5 C-terminal amino acids of G<sub>p</sub>a1, i.e., positions 468-472, are removed and replaced with the last 5 C-terminal amino acids of a heterologous G  $\alpha$  subunit. The C-terminal amino acids of the G<sub>p</sub>a1 amino acids can be replaced with C-terminal amino acids of essentially any heterologous G protein  $\alpha$  subunit, but most preferably are replaced with mammalian sequences, most preferably mouse, rat or human sequences. The heterologous G  $\alpha$  subunits can be selected from any member of the G  $\alpha$  subunit family, for example G<sub>a</sub>s, G<sub>a</sub>i2, G<sub>a</sub>i3, G<sub>a</sub>q, G<sub>a</sub>o<sub>a</sub>, G<sub>a</sub>o<sub>b</sub> and G<sub>a</sub>16. Preferably, the G  $\alpha$  subunit is selected from G<sub>a</sub>s, G<sub>a</sub>i2, G<sub>a</sub>i3, G<sub>a</sub>q, and G<sub>a</sub>16.

In one embodiment, from the last 4 to the last 50 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In another embodiment, from the last 4 to the last 40 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In another embodiment, from the 5 last 4 to the last 28 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In a preferred embodiment, from the last 4 to the last 6 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In a more preferred embodiment, the last 5 C-terminal amino acids of the heterologous G  $\alpha$  subunit replace the C-terminal amino acids of G $\alpha$ 1. In various other 10 embodiments, any number of C-terminal amino acid residues of the heterologous G  $\alpha$  subunit from 4 to 50 amino acids can replace G $\alpha$ 1 sequences, e.g., 4-50, 4-40, 4-30, 4-28, 4-6, more preferably, 5 or 6.

Additionally, for sandwich chimera, N-terminal amino acids of a heterologous G  $\alpha$  subunit replace or are added onto N-terminal amino acids of G $\alpha$ 1. Thus, the N- 15 terminus of the G $\alpha$ 1 is "operably linked" to the N-terminal amino acids of a heterologous G  $\alpha$  subunit. The N-terminal amino acids can be derived from the same heterologous G  $\alpha$  subunit as the C-terminal sequences or from a different heterologous G  $\alpha$  subunit. Most preferably, the heterologous N-terminal amino acids are from a mammalian G  $\alpha$  subunit, most preferably mouse, rat or human sequences. The 20 heterologous G  $\alpha$  subunits can be selected from any member of the G  $\alpha$  subunit family, for example G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3, G $\alpha$ q, G $\alpha$ o<sub>a</sub>, G $\alpha$ o<sub>b</sub> and G $\alpha$ 16. Preferably, the G  $\alpha$  subunit is selected from G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3, G $\alpha$ q, and G $\alpha$ 16.

The N-terminal amino acids of G $\alpha$ 1 can be deleted and replaced with N-terminal amino acids of a heterologous (e.g. mammalian) G  $\alpha$  subunit, for example, N-terminal amino acids 1-50 of G $\alpha$ 1, can be removed and replaced with N-terminal amino acids from the heterologous G  $\alpha$  subunit. In another embodiment, N-terminal amino acids 1-30 of G $\alpha$ 1 can be removed and replaced with N-terminal amino acids from the heterologous G  $\alpha$  subunit. In a preferred embodiment, N-terminal amino acids 1-21 of the G $\alpha$ 1 are removed and replaced with N-terminal amino acids from the heterologous 30 G  $\alpha$  subunit. In another preferred embodiment, N-terminal amino acids 1-11 of G $\alpha$ 1 are removed and replaced with N-terminal amino acids from the heterologous G  $\alpha$  subunit.

In one embodiment, the first 50 N-terminal amino acids of the heterologous G  $\alpha$  subunit replace or are added onto the N-terminal amino acids of G $\alpha$ 1. In another embodiment, the first 40 N-terminal amino acids of the heterologous G  $\alpha$  subunit 35

replace or are added onto the N-terminal amino acids of G<sub>p</sub>a1. In another embodiment, the first 30 N-terminal amino acids of the heterologous G  $\alpha$  protein subunit replace or are added onto the N-terminal amino acids of G<sub>p</sub>a1. In another embodiment, the first 21 N-terminal amino acids of the heterologous G  $\alpha$  protein subunit replace or are added 5 onto the N-terminal amino acids of G<sub>p</sub>a1. In yet another embodiment, the first 11 N-terminal amino acids of the heterologous G  $\alpha$  protein subunit replace or are added onto the N-terminal amino acids of G<sub>p</sub>a1. In various other embodiments, any number of N-terminal amino acid residues of the heterologous G  $\alpha$  subunit from 5 to 50 amino acids can replace or be added onto N-terminal G<sub>p</sub>a1 sequences, e.g., 5-50, 5-40, 5-30, 5-21 or 10 5-11.

In one embodiment, replacements to the C-terminal and N-terminal amino acids of G<sub>p</sub>a1 can be made using the same heterologous G  $\alpha$  subunit. In one embodiment, this heterologous G  $\alpha$  subunit is the G $\alpha$ q. The skilled artisan will appreciate that replacements can be made with any member of the G  $\alpha$  subunit family. Also within the 15 scope of the invention, the replacements at the N- and C-termini can be made with different G  $\alpha$  subunits at the C-terminal and N-termini of G<sub>p</sub>a1. Examples of combinations of C-terminal and N-terminal replacements include, but are not limited to, G $\alpha$ s-G<sub>p</sub>a1-G $\alpha$ q; G $\alpha$ q-G<sub>p</sub>a1-G $\alpha$ s; G $\alpha$ s-G<sub>p</sub>a1-G $\alpha$ i2; G $\alpha$ i2-G<sub>p</sub>a1-G $\alpha$ s; G $\alpha$ s-G<sub>p</sub>a1-G $\alpha$ i3; G $\alpha$ i3 -G<sub>p</sub>a1-G $\alpha$ s, and the like.

20 The G<sub>p</sub>a1 sequence preferably is replaced with a wild type sequence of a heterologous G  $\alpha$  subunit, for example by replacing the last 5 C-terminal amino acids of the G<sub>p</sub>a1 with the corresponding last 5 C-terminal amino acids of a wildtype mammalian G  $\alpha$  subunit. However, the G<sub>p</sub>a1 sequence may also be replaced with sequences of a mutant G  $\alpha$  subunit containing, for example, one or more conservative 25 substitutions in the last 5 C-terminal amino acid sequence of the G  $\alpha$  subunit. Mutant G  $\alpha$  subunit sequences are selected based on their ability to retain, or improve, receptor specificity. Accordingly, the C-terminal amino acids of the heterologous G $\alpha$  subunit are intended to include wildtype sequences and conservative substitutions that maintain receptor specificity.

30 Coupling specificity of the chimeric G $\alpha$  proteins can be tested by expressing the sandwich chimeric G proteins in yeast cells that also express a heterologous G protein-coupled receptor. Coupling of the sandwich chimeric G proteins to a heterologous receptors can be determined by, for example, Fus1-pHIS3 screening, as described in the Example 12. Sandwich chimeric G proteins in which C-terminal and N-terminal amino 35 acids of G<sub>p</sub>a1 are replaced by sequences of heterologous G  $\alpha$  subunits demonstrate an

improved coupling when tested with various heterologous G-protein coupled receptors (see Example 12).

All patents, published patent applications and other references disclosed  
5 herein are hereby expressly incorporated by reference.

### EXEMPLIFICATION

The invention now being generally described will be more readily understood by reference to the following examples, which are included merely for  
10 purposes of illustration of certain aspects and embodiments of the present invention and are not intended to limit the invention.

#### **Example 1: Construction of chimeric yeast/mammalian G $\alpha$ subunits.**

This Example pertains to the construction of chimeric G protein subunits which are made using a first polypeptide derived from a yeast G protein subunit and a second polypeptide derived from a mammalian G protein subunit. The G $\alpha$  subunit of heterotrimeric G proteins interacts with both the  $\beta\gamma$  complex and the receptor. The construction of chimeric  $\alpha$  subunits can, therefore, enhance receptor coupling. In this  
15 example, various G protein subunit chimeras were constructed using a polypeptide derived from yeast GPA1 and several different mammalian G $\alpha$  subunits. The nucleotide sequence of GPA1 is known in the art. For example, the gene encoding a G protein homolog of *S. cerevisiae* was cloned independently by Dietzel and Kurjan (supra) (who referred to the gene as SCG1) and by Nakafuku, et al. (1987 Proc Natl Acad Sci  
20 84:2140-2144) (who called the gene GPA1). The nucleotide sequence of the human G alpha subunits G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3 also are available in the art. The following information provides specific examples of the construction of G protein subunit, G $\alpha$ , chimeras.  
25

The plasmids used were constructed as follows.

30 **pRS416-GPA1 (Cadus 1069).** An XbaI - SacI fragment encoding the entire GPA1 promoter region, coding region and approximately 250 nucleotides of 3' untranslated region was excised from 10 YCplac111-GPA1 (from S. Reed, Scripps Institute) and cloned into YEp vector pRS416 (Sikorski and Hieter, Genetics 122: 19 (1989)) cut with XbaI and SacI.

- 74 -

**Site-directed mutagenesis of GPA1 (Cadus 1075, 1121 and 1122).** A 1.9 kb EcoRI fragment containing the entire GPA1 coding region and 200 nucleotides from the 5' untranslated region was cloned into EcoRI cut, phosphatase-treated pALTER-1 (Promega) and transformed by electroporation (Biorad Gene Pulser) into 5 DH5 $\alpha$ F' bacteria to yield Cadus 1075. Recombinant phagemids were rescued with M13KO7 helper phage and single stranded recombinant DNA was extracted and purified according to the manufacturer's specifications. A new NcoI site was introduced at the initiator methionine of GPA1 by oligonucleotide directed mutagenesis using the synthetic oligonucleotide:

10

5' GATATATTAGGTAGGAAACCATGGGTGTACAGTGAG 3'. (SEQ ID NO: 13).

15

Positive clones were selected in ampicillin and several independent clones were sequenced in both directions across the new NcoI site at +1. Two clones containing the correct sequences were retained as Cadus 1121 and 1122.

20

**Construction of a GPA1-based expression vector (Cadus 1127).** The vector used for expression of full length and hybrid mammalian Go proteins in yeast, Cadus 1127, was constructed in the following manner. A 350 nucleotide fragment spanning the 3' untranslated region of GPA1 was amplified with Taq polymerase (AmpliTaq; Perkin Elmer) using the oligonucleotide primers A: 5' CGAGGCTCGAGGGAACGTATAATTAAAGTAGTG 3' (SEQ ID NO: 14) and B: 5' GCGCGGTACCAAGCTTCAATTGAGATAATACCC 3' (SEQ ID NO: 15). The 350 nucleotide product was purified by gel electrophoresis using GeneClean II (Bio101) and 25 was cloned directly into the pCRII vector by single nucleotide overlap TA cloning (InVitrogen). Recombinant clones were characterized by restriction enzyme mapping and by dideoxynucleotide sequencing. Recombinant clones contained a novel XbaI site 5' to the authentic GPA1 sequence and a novel KpnI site 3' to the authentic GPA1 sequence donated respectively by primer A and primer B.

30

The NotI and SacI sites in the polylinker of Cadus 1013 (pRS414) were removed by restriction with these enzymes followed by filling in with the Klenow fragment of DNA polymerase I and blunt end ligation to yield Cadus 1092. The 1.4 kb PstI - EcoRI 5' fragment of GPA1 from YCplac111-GPA1 containing the GPA1 35 promoter and 5' untranslated region of GPA1 was purified by gel electrophoresis using GeneClean (Bio101) and cloned into PstI - EcoRI restricted Cadus 1013 to yield Cadus

- 75 -

1087. The PCR amplified XhoI - KpnI fragment encoding the 3' untranslated region of  
GPA1 was excised from Cadus 1089 and cloned into XhoI - KpnI restricted Cadus 1087  
to yield Cadus 1092. The NotI and SacI sites in the polylinker of Cadus 1092 were  
removed by restriction with these enzymes, filling in with the Klenow fragment of DNA  
5 polymerase I, and blunt end ligation to yield Cadus 1110. The region of Cadus 1122  
encoding the region of GPA1 from the EcoRI site at -200 to +120 was amplified with  
Vent DNA polymerase (New England Biolabs, Beverly, MA) with the primers  
5' CCCGAATCCACCAATTCTTTACG 3' (SEQ ID NO: 16) and  
5' GCGGCCTCGACGGGCCGCGTAACAGT 3' (SEQ ID NO: 17).

10 The amplified product, bearing an EcoRI site at its 5' end and novel SacI,  
NotI and SalI sites at its 3' end was restricted with EcoRI and SalI, gel purified using  
GeneClean II (Biol01), and cloned into EcoRI and SalI restricted Cadus 1110 to yield  
Cadus 1127. The DNA sequence of the vector between the EcoRI site at -200 and the  
KpnI site at the 3' end of the 3' untranslated region was verified by restriction enzyme  
15 mapping and dideoxynucleotide DNA sequence analysis.

**PCR amplification of GPA<sub>41</sub>-G $\alpha$  proteins and cloning into Cadus  
1127.** cDNA clones encoding the human G alpha subunits G $\alpha$ s, G $\alpha$ i2, G $\alpha$ i3, and S.  
cerevisiae GPA1 were amplified with Vent thermostable polymerase (New England  
20 Biolabs, Beverly, MA). The primer pairs used in the amplification are as follows:

For G $\alpha$ S: Primer 1 : 5' CTGCTGGAGCTCCGCCTGCTGCTGGGTGCTGGAG 3'  
(SEQ ID NO: 18) (SacI 5')

Primer 2 :

25 5' CTGCTGGTCGACGCCGCGGGGGTTCTTCTTAGAAGCAGC 3' (SEQ ID NO:  
19) (SalI 3')

Primer 3 : 5' GGGCTCGAGCCTTCTTAGAGCAGCTCGTAC 3'  
(SEQ ID NO: 20) (XhoI 3')

30 For G $\alpha$ i2: Primer 1 : 5' CTGCTGGAGCTCAAGTTGCTGCTGTTGGGTGCTGGGG 3'  
(SEQ ID NO: 21) (SacI 5')

Primer 2 :

5' CTGCTGGTCGACGCCGCGCCCTCAGAAGAGGCCGCGGT CC 3' (SEQ ID NO:  
22) (SalI 3')

35 Primer 3 : 5' GGGCTCGAGCCTCAGAAGAGGCCGCGACTC 3'  
(SEQ ID NO: 23) (XhoI 3')

- 76 -

for G<sub>αi3</sub>: Primer 1: 5' CTGCTGGAGCTCAAGCTGCTGCTACTCGGTGCTGGAG3' (SEQ ID NO: 24) (SacI 5')

Primer 2:

5 5' CTGCTGGTCGACGCCGGCCACTAACATCCATGCTTCTCAAT AAAGTC3' (SEQ ID NO: 25) (SalI 3')

Primer 3 : 5' GGGCTCGAGCATGCTTCTCAATAAAGTCCAC3' (SEQ ID NO: 26) (XhoI 3')

10 After amplification, products were purified by gel electrophoresis using GeneClean II (Bio101) and were cleaved with the appropriate restriction enzymes for cloning into Cadus 1127.

The hybrid GPA<sub>41</sub>-G<sub>α</sub> subunits were cloned via a SacI site introduced at the desired position near the 5' end of the amplified genes and a SalI or XhoI site 15 introduced in the 3' untranslated region. Ligation mixtures were electroporated into competent bacteria and plasmid DNA was prepared from 50 cultures of ampicillin resistant bacteria.

**Construction of Integrating Vectors Encoding GPA<sub>41</sub>-G<sub>α</sub> Subunits.**

20 The coding region of each GPA<sub>41</sub>-G<sub>α</sub> hybrid was cloned into an integrating vector (pRS406 = URA3 AmpR) using the BssHII sites flanking the polylinker cloning sites in this plasmid. Cadus 1011 (pRS406) was restricted with BssHII, treated with shrimp alkaline phosphatase as per the manufacturer's specifications, and the linearized vector 25 was purified by gel electrophoresis. Inserts from each of the GPA<sub>41</sub>-G<sub>α</sub> hybrids were excised with BssHII from the parental plasmid, and subcloned into gel purified Cadus 1011.

**Construction of GPA<sub>BAM</sub>-G<sub>α</sub> Constructs.** A novel BamHI site was introduced in frame into the GPA1 coding region by PCR amplification using Cadus 30 1179 (encoding a wildtype GPA1 allele with a novel NcoI site at the initiator methionine) as the template, VENT polymerase, and the following primers: Primer A = 5' GCATCCATCAATAATCCAG 3' (SEQ ID NO: 27) and Primer B = 5' GAAACAATGGATCCACTTCTTAC 3' (SEQ ID NO: 28). The 1.1 kb PCR product was gel purified with GeneClean II (Bio101), restricted with NcoI and BamHI and cloned 35 into NcoI-BamHI cut and phosphatased Cadus 1122 to yield Cadus 1605. The sequence of Cadus 1605 was verified by restriction analysis and dideoxy-sequencing of double-

stranded templates. Recombinant  $\text{GPA}_{\text{Bam}}$ - $\text{G}\alpha$  hybrids of  $\text{G}\alpha s$ ,  $\text{G}\alpha i2$ , and  $\text{G}\alpha 16$  were generated. Construction of Cadus 1855 encoding recombinant  $\text{GPA}_{\text{Bam}}$ - $\text{G}\alpha 16$  serves as a master example: construction of the other hybrids followed an analogous cloning strategy. The parental plasmid Cadus 1617, encoding native  $\text{G}\alpha 16$ , was restricted with NcoI and BamHI, treated with shrimp alkaline phosphatase as per the manufacturer's specifications and the linearized vector was purified by gel electrophoresis. Cadus 1605 was restricted with NcoI and BamHI and the 1.1 kb fragment encoding the amino terminal 60% of GPA1 with a novel BamHI site at the 3' end was cloned into the NcoI- and BamHI-restricted Cadus 1617. The resulting plasmid encoding the  $\text{GPA}_{\text{Bam}}$ - $\text{G}\alpha 16$  hybrid was verified by restriction analysis and assayed in tester strains for an ability to couple to yeast  $\text{G}\beta\gamma$  and thereby suppress the  $\text{gpa}1$  null phenotype. Two additional  $\text{GPA}_{\text{Bam}}$ - $\text{G}\alpha$  hybrids,  $\text{GPA}_{\text{Bam}}$ - $\text{G}\alpha s$  and  $\text{GPA}_{\text{Bam}}$ - $\text{G}\alpha i2$ , described in this application were prepared in an analogous manner using Cadus1606 as the parental plasmid for the construction of the  $\text{GPA}_{\text{Bam}}$ - $\text{G}\alpha i2$  hybrid and Cadus 1181 as the parental plasmid for the construction of the  $\text{GPA}_{\text{Bam}}$ - $\text{G}\alpha s$  hybrid.

**Example 2. Identification of critical regions of  $\text{G}\alpha$  and improvement of the interaction of mammalian  $\text{G}\alpha s$  with yeast  $\beta\gamma$  by mutation.**

In this Example, critical regions of various mammalian  $\text{G}\alpha$  subunits were identified by making the following series of mutations:

D229S in rat  $\text{G}\alpha s$ , S270P in human  $\text{G}\alpha 16$ , S280P in  $\text{G}\alpha 1-\text{G}\alpha i2-\text{G}\alpha oB$  and S288P in  $\text{GPA}1-\text{G}\alpha 12$

The specificity of coupling of a receptor to a heterotrimeric G-protein is largely determined by the  $\alpha$  subunit of the G-protein. This fact has been exploited by expressing human 7 transmembrane receptors in yeast and functionally coupling them into the yeast pheromone response pathway, as a path to drug discovery. In order to accomplish this goal mammalian  $\text{G}\alpha$  subunits were expressed in yeast cells whose own pheromone-responsive  $\text{G}\alpha$  subunit (GPA1), has been disrupted. As the predominant role of GPA1 is to bind to and sequester the effector-signaling  $\beta\gamma$  component of the heterotrimer, the effector activity of the mammalian  $\text{G}\alpha$  is irrelevant; the activity sought is the ability to bind yeast  $\beta\gamma$  in the quiescent state, and to release it upon receptor activation.

In general, mammalian  $\text{G}\alpha$  subunits couple poorly to the  $\beta\gamma$  subunits of  $\text{Saccharomyces cerevisiae}$ . When these  $\text{G}\alpha$ 's are expressed in yeast that lack their own endogenous  $\text{G}\alpha$  subunit, this failure to couple results in the constitutive activation of the

- 78 -

pheromone pathway, due to the effector activity of the unbound yeast  $\beta\gamma$ . The interaction of mammalian G $\alpha$ s with yeast  $\beta\gamma$  can be improved through a variety of random and semirandom mutagenesis techniques in conjunction with various genetic selection and screening methods. Such an analysis has highlighted critical regions for 5  $\alpha$ - $\beta\gamma$  coupling, as well as identifying residues in G $\alpha$  that do not directly bind to  $\beta\gamma$  but that nevertheless mediate this coupling.

### Yeast Strains

The following yeast strains were used in these experiments:

10 CY7757 (MAT $\alpha$  stel 4::trpl ::LYS2 gpal ::LEU2 fus1-HIS3 far1-1 trpl his3 ura3 leu2 lys2 ade2-1), and CY1316 ( MAT $\alpha$  gpa $\Delta$ 1163 far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 stel4::trpl::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trpl his3) were constructed using standard genetic techniques. The ste14::TRP1, gpa $\Delta$ 1163 alleles were derived from GMZ809 (Hrycyna et al. 1991. EMBO J. 10:1699), D111 15 (Kurjan et al. 1987. Cell 50:1001), IH2512 (Chang et al. 1990. 63:999), and SY1390 (Stevenson et al. 1992. Genes and Development 6:1293), respectively. The gpa $\Delta$ 1163 (a deletion of the internal SphI fragment of GPA1), ste3 $\Delta$ 1156 (a deletion extending from the start codon to the stop codon of STE3), and far1 $\Delta$ 1442 (a deletion from codon 50 through 696 of FAR1) alleles were introduced by two-step gene disruption using 5- 20 fluoroorotic acid (Rothstein. 1991. Methods in Enzymol. Vol. 194). The ste14::trpl::LYS2 lesion was constructed using pRS624 (Sikorski and Hieter. 1989. Genetics).

### Plasmids

25 For expression of the various G $\alpha$  constructs, an expression plasmid derived from the centromere plasmid C p1127 was used. This plasmid contains the TRP1 gene as a selectable marker and the GPA1 promoter and terminator separated by an NcoI and XhoI site that act as 5' and 3' cloning sites respectively. Cp3098 contains the rat G $\alpha$ s subunit subcloned from pYSK136 (Dietzel and Kurjan, 1987. Cell. 50:1001), using the 5' NcoI site contained therein and a 3' XhoI site introduced by PCR using the oligonucleotide "gsxhorev" (5' - CCCCTCGAGTTCCCTTCTAGAGCAGCT) (SEQ ID NO:45). Cp3222 was constructed by replacing the 917 bp Hind III fragment containing the TRP1 gene with a 2.4 kb BbsI-digested PCR fragment encoding the ADE2 gene amplified from genomic wild type yeast DNA using the oligonucleotides 30 "ADE2FWD" (5'-CCCGAAGACCAAGCTTTGACCAGGTTATTATA) (SEQ ID NO:46) and "ADE2REV" (5' -AAGGAAGACTTAGCTTATAATTGGGCTTAGTT) (SEQ ID 35

NO:47); the ADE2 gene in this plasmid is transcribed in the opposite direction from the Gas cassette.

The mutation D229S in the wild type rat Gαs was made by amplifying two fragments from the rat Gαs template Cp3098, using the oligo pairs 523 (TTT CTT GTC ACT CCG TTT CTA AC) (SEQ ID NO:48) and D229S reverse (CCCCGTCTCAAGAGCGCTGGCCGCCACATC) (SEQ ID NO:49); and D229S forward (CCCCGTCTCACTCTGAACGCCGCAAGTGGATCC) (SEQ ID NO:50) and 759 (AGC AAG CAG ATC TTG CTT GTT G) (SEQ ID NO:51). The fragment resulting from the former pair was digested with NcoI and BsmBI, and the fragment from the latter pair was digested with BglII and BsmBI, and together they were ligated into Cp3098 digested with NcoI and BglII to create Cp3390. This plasmid harbored the sequence:

15 GAT GTG GGC GGC CAG CGC TCT GAA CGC CGC  
           AAG TGG ATC (SEQ ID NO:52)  
 D V G G Q R S E R R K  
     W I  
 (SEQ ID NO:53)

20 Thereby encoding rat Gas having a substitution of serine for the wild-type aspartic acid  
at position 229.

The plasmid Cp3699, encoding a G<sub>pal</sub>-G<sub>α16</sub> chimera with the S270P mutation, was constructed as follows. KS5714

25 (SEQ ID NO:54) and KS\_1661

(SEQ ID NO:54) and 5'-CTCGCTGGTCGACGCCGCGGGTCACAGCAGGTTGAT CTCGTCCAG3' (SEQ ID NO:55) were used to amplify the human Gal6 gene from the 45th to the 374th codon, and this fragment was digested with SacI and SalI, ligated to Cp1127 digested

30 chimera. This plasmid was mutated (S270P) by amplification with the primer pairs G16-  
 271F (GCTAGGTCTCAACCAGTCATCCTCTTCTAACAAAACC) (SEQ ID NO:56)  
 and 524 (ACC CGG AAC GAT TTA ACG AG) (SEQ ID NO: 84), and G16-271R  
 (ACGTGGTCTCATGGTGTGCTTTGAACCAGGGTAGT) (SEQ ID NO:57) and 713  
 (AGC GGC TGC AGA TTC CAT TC) (SEQ ID NO:58), digestion of the resulting  
 35 fragments with Bsal and Xhol, or with Bsal and NheI, respectively, and ligation together

- 80 -

into Cp3233 digested with NheI and Xhol to yield Cp3699, harboring the following mutated sequence:

5      CTA   CCC   TGG   TTC   AAA   AGC   ACA   **CCA**   GTC   ATC   CTC   TTT  
 CTC   AAC   AAA  
 (SEQ ID NO:59)  
 L   P   W   F   K   S   T   P   V   I   L   F  
 L   N   K  
 (SEQ ID NO:60)

10 thereby encoding human G $\alpha$ 16 having a substitution of proline for the wild-type serine at position 270.

Plasmid Cp3635, encoding a G $\alpha$ l-G $\alpha$ i2-G $\alpha$ oB chimera with the S280P mutation, was constructed as follows. G $\alpha$ oB was amplified from Cp 1131 using the primers ob forward (GGGCGTCTCACATGGGATGTACGCTGAGCG) (SEQ ID NO:61) and ob reverse (GGGGTCGACTCAGTAGAGTCAACAGCCC) (SEQ ID NO:62), and the resulting fragment was digested with BsmBI and SalI, and ligated into Cpl127 digested with NcoI and SalI, yielding Cp3332. The S280P mutation was introduced by amplification with the primer pairs 565 (GAT TGG AGC CGG TGA CTA CC) (SEQ ID NO: 85) and obsp reverse

20 (CCCCCGTCTCATAGGGTATCAATGAAAAACTTGTGTTA) (SEQ ID NO:63), and 524 (supra.) and obsp forward (CCCCCGTCTCACCTATCATCCTCTTCAACAAG) (SEQ ID NO:64), digestion of the resulting fragments with AatII and BsmBI, or with BsmBI and Sal I, respectively, and ligation together into Cp3332 digested with AatII and SalI to yield Cp3709, harboring the following mutated sequence:

25      AAC   AAC   AAG   TTT   TTC   ATT   GAT   ACC   CCT   ATC   ATC   CTC   TTC   CTC   AAC  
 AAG  
 (SEQ ID NO:65)  
 N   N   K   F   F   I   D   T   P   I   I   L   F   L   N  
 30   K  
 (SEQ ID NO:66)

The italicized nucleotides and amino acids indicated above refer to polymorphisms derived from G $\alpha$ A. (The corresponding G $\alpha$ oB amino acids are W and T rather than F and I.) An SphI-Xhol fragment from Cp3709 was ligated into Cpl 183 digested with the

35

same enzymes, to yield Cp3635, encoding a Gp1(1-43)-G $\alpha$ 12(36-242)-G $\alpha$ B(243-354) chimera containing the above changes.

Plasmids encoding GPA1-G $\alpha$ 12 chimerae were constructed as follows.

- 5 The single Pst I site in Cp1127 was eliminated by digesting with BamH1 and PstI, blunting the overhangs with T4 DNA polymerase, and ligating the resulting linear DNA; the resulting plasmid (Cp3326) was digested with NcoI and SacI, allowing insertion of a synthetic oligonucleotide encoding the N-terminus 41 amino acids of GPA1p. The synthetic oligonucleotide, which contains a Pst I site at codons 18 and 19 of GAP1,  
10 was made by annealing phosphorylated oligonucleotide o207  
(AAAAGAGCCAATGATGTCATCGAGCAATCGTTGCAGCTGGAGAAACAAACGTGACAAGA ATGAGCT) (SEQ ID NO:67) with oligonucleotide o208  
(CATTCTTGTACAGTTGTTCTCCAGCTGCAACGATTGCTCGATGACATCATTGGCTCT TTTGTTC) (SEQ ID NO:68) and oligonucleotide o209  
15 (CATGGGGTGTACAGTGAGTACGCAAACAATAGGAGATGAAAGTGATCCTTCTGCAG AAC) (SEQ ID NO:69) with phosphorylated oligonucleotide o210  
(TGCAGAAAAGGATCACTTCATCTCCTATTGTTGCGTACTCACTGTACACCC) (SEQ ID NO:70), followed by their ligation. The plasmid resulting from the insertion of this approximately 120 bp synthetic DNA fragment, Cp3363, was then digested with  
20 SacI and XhoI, permitting the insertion of a PCR-amplified, SacI- and XhoI-digested fragment encoding amino acids 43 through the stop codon of wild-type G $\alpha$ 12. The resulting plasmid Cp3435 thus encodes a chimeric GPA1-G $\alpha$ 12 protein in which the N-terminal 41 residues of GPA1 replace the N-terminal 42 amino acids of wild-type G $\alpha$ 12. An equivalent PCR-amplified, SacI- and XhoI-digested fragment encoding amino acids  
25 43 through the stop codon of the GTPase-deficient mutant of G $\alpha$ 12 was also inserted into Cp3363. The resulting plasmid Cp3436 encodes a chimeric protein in which the N-terminal 41 residues of GPA1 replace the N-terminal 42 amino acids of the Q229L mutant of G $\alpha$ 12.

- Cp3435 was used as the template for mutagenesis by Stratagene's Quik-  
30 Change protocol using oligonucleotides o286 (CTTCTTCAACGTCCCCATCATCCTC) (SEQ ID NO:71) and o287 (GAGGATGATGGGGACGTTGAAGAAG) (SEQ ID NO:72) to create Cp3822. Cp3822 encodes a GPA1-G $\alpha$ 12 chimera in which the serine corresponding to residue 288 of wild-type G $\alpha$ 12 has been changed to a proline. Cp3435 was also the template for mutating the glycine corresponding to residue 228 of wild-type  
35 G $\alpha$ 12 to alanine using oligonucleotides o293 (GGATGTGGCGCCAGAGGTCACAG) (SEQ ID NO:73) and o294

- 82 -

(CTGTGACCTCTGGCGCCACATCC) (SEQ ID NO:74). This mutant encodes a chimeric G $\alpha$  subunit that is not likely capable of assuming an activated conformation (E. Lee, R. Taussig, and A.G. Gilman, J. Biol. Chem. 267, 1212 (1992); R.T. Miller, et al., Nature 334, 712 (1988)).

5 Finally, a 1.9 kb fragment containing the ADE8 gene was cloned into the Hpa I site to Cp3435 to create Cp3506. The unique Pst I and SacI sites of Cp3506 were then used to introduce oligonucleotide libraries that encode variable amino acids. That is, Cp3506 was digested with PstI and SacI for the insertion of a library of semi-random oligonucleotides constructed as follows: oligonucleotides o221  
10 (GCGGAGCTCMNNMNNMNNMNNCTTTCTAATTGCAAGGATTGTTCGATAACGTCA TTAGCTCTTTATTCTGCAGGG, where M is C or A and N is A, G, C, or T) (SEQ ID NO:75) and o222  
(CCCTGCAGAATAAGAGAGCTAATGACGTTATCGAACAACTTGCATT) (SEQ ID NO:76) were annealed. the partially double-stranded oligonucleotides were filled in with  
15 Sequenase, and the fully double-stranded product was cut with SacI and PstI. The resulting library encodes G $\alpha$  chimerae with the following sequence: The N-terminal 35 amino acids are those of GPA1, followed in succession by 5 completely random amino acids in place of residues 36-40 of GPA1, residues 41-44 of GPA1, and residues 59-380 (stop) of G $\alpha$ 12. The random region corresponds to the C-terminal border of the  
20 predicted N-terminal  $\alpha$  helix. This library was used as a possible source of chimerae that may exhibit improved coupling with STE4/STE18 as a result of optimal positioning of the N-terminal  $\alpha$  helix relative to the rest of the G $\alpha$  subunit.

Wild-type G $\alpha$ 12 shows no greater ability to couple to STE4/STE18 than  
25 empty vector. GPA1-G $\alpha$ 12, however, shows clearly improved coupling when assayed at 2.5 mM aminotriazole, with barely detectable improvement at 1 mM aminotriazole. Thus, replacement of the N-terminal G $\alpha$  helix of G $\alpha$ 12 with the equivalent region of GPA1 increases the coupling of the G $\alpha$  subunit with STE4/STE18. Even greater coupling is seen with the GPA1-G $\alpha$ 12 S288P mutant: coupling is clear at 0.5 mM AT.  
30 In contrast, neither the likely GTPase-deficient GPA1-G $\alpha$ 12 Q227L chimera nor the GPA1-G $\alpha$ 12 G228A chimera show greater coupling than wild-type G $\alpha$ 12.

The library of GPA1-G $\alpha$ 12 chimerae that vary in amino acid sequence at the C-terminal border of the predicted N-terminal  $\alpha$  helix was screened for proteins that show improved coupling to STE4/STE18. No such proteins were found.

**Media.** For the propagation of yeast cells harboring chimeric G protein subunits, standard recipes were used for yeast media. SCA and contains a mixture of amino acids and uracil (See e.g., Rose et al. Methods in Yeast Genetics, 1990), and lacks adenine; the pH was adjusted to 6.8 using 4.5 ml 1M KOH and 25 ml 1M K-Pipes pH6.8 per liter.

- 5 SCAH1 is the same except it lacks histidine and contains lmM (1,2,4)3-aminotriazole, a competitive inhibitor of the His3 gene product. SCH1(loA) is identical to SCAH1 except that it contains 6.25 µg/ml adenine.

**Mutant Libraries.** The rat G $\alpha$ s protein was mutagenized by taking advantage of the error-prone nature of Taq polymerase in the presence of dITP. Rat G $\alpha$ s was amplified from p3098 for 30 cycles (30 sec 94; 30 sec 50; 90 sec 72) using 705 (5'-GCATCACATCAATAATCCAG) (SEQ ID NO:77) and 386 (5'-AACCCGGAACGATTAAACGAGATCAAGAAC) (SEQ ID NO:78) as primers (these oligonucleotides correspond to the GPAl promoter and terminator, respectively) and 200 15 µM dATP, dGTP, dTTP and dITP, and 40pM dCTP, and the buffer supplied by the manufacturer (Fisher). The PCR product was digested with NcoI and XhoI and ligated into the ADE2 vector described above. Approximately 20,000 DH10B transformants were pooled to create a library of rat G $\alpha$ s mutants.

20 **Screening results**

The screen for G $\alpha$ s mutants that show improved binding to yeast G $\beta$  $\gamma$  takes advantage of the fact that a gpal fus1 -HIS3 colony expressing wild type G $\alpha$ s can grow upon replica plating to media lacking histidine and containing lmM 3-AT, due to the partially constitutive state of the pheromone pathway, which leads to partial 25 derepression of the fus1-HIS3 reporter gene. A colony containing the desired  $\beta$  $\gamma$ -coupling G $\alpha$ s mutant will presumably fail to grow on this media due to the inactivation of the pheromone pathway mediated by tight  $\alpha$  $\beta$  $\gamma$  association. However, if the selection for the G $\alpha$ s-bearing plasmid is relaxed, in this case by the addition of limiting amounts of adenine, then the desired mutant colony will only produce growth from that fraction of 30 cells in the colony (usually 5-10%) that have lost the G $\alpha$ s plasmid. Such a colony will appear red due to the buildup of an intermediate of the adenine pathway. The nonmutated G $\alpha$ s-containing colonies will appear white on an adenine-limiting plate, since the presence of the G $\alpha$ s-ADE2 plasmid has no appreciable negative effect on the ability of the cells to grow in lmM 3-AT, but confers a selective advantage under 35 adenine-limiting conditions. The visual aspect of the screen allows the facile

- 84 -

identification of potential mutants, and eliminates potential unwanted mutations, such as Ste-, because the screen demands a plasmid-dependent phenotype.

A library of mutagenized rat G<sub>α</sub>s genes under the control of the GPA1 promoter was introduced into a gp<sub>1</sub>fus1-HIS3 far1 ade2 strain (CY7757) and approximately 10,000 colonies were selected on 40 SCA plates. After 2 days of growth at 30°, the colonies were replica plated to SCH1(loA) plates to identify those bearing mutant G<sub>α</sub>s proteins capable of suppressing the pheromone pathway through a tighter interaction with endogenous yeast βγ; such colonies appear red due to the outgrowth of Ade- cells, because the Ade+ component of the colony is incapable of overcoming the selection for His+, which demands a pheromone pathway at least partially derepressed. As expected, colonies that produce red replicas on SCH1(loA) are unable to grow on SCAH1, where the ADE2-G<sub>α</sub>s plasmid is strictly required.

Plasmid DNA was recovered from putative mutant-bearing cells, amplified in E. coli DH10B, retransformed into CY7757 and retested on SCAH1 to confirm the plasmid linkage of the mutation. Sequencing of the mutants revealed single or double point mutations as described in Table 1 below. In all cases, subcloning of the mutations into a wild type G<sub>α</sub>s-TRP1 plasmid (p3098) as performed to rule out the possibility that mutations in noncoding sequences confer the His<sup>+</sup>phenotype. These subcloned plasmids were tested in a different gp<sub>1</sub>fus1-HIS3 strain, CY1316.

**Table 1.**

<u>Mutant</u>	<u>Mutations</u>
rN5-4; hN7; hN8-1; rI7-5	<b>E10K</b> (GAG→AAG)
hI-8	<b>D229V</b> (GAT→GTT); <b>S82N</b> (AGC→AAC)
hN8-2	<b>S286P</b> (TCT→CCT); <b>E75V</b> (GAG→GTG)
rA	<b>S286P</b> (TCT→CCT)
rJ	<b>N254D</b> (AAC→GAC)

For those mutants with two mutations, the bold type indicates the mutation that confers the improved coupling to G<sub>α</sub>s, as determined by subcloning into Cp3098.

Figure 2 depicts the behavior of the G $\alpha$ s mutations. All except S286P eliminate the growth of the gpal fusl-HIS3 strain on 0.2mM 3-AT, indicating efficient suppression of the pheromone pathway. This effect may result from an improved affinity of G $\alpha$ s-GDP for yeast G $\beta$  $\gamma$ , without discounting other possible explanations, such as a reduction in  
5 GTP affinity or an augmented GTPase activity.

Examination of the alignments between various families of G $\alpha$  subunits revealed that despite the high degree of conservation in the "Switch 2" region among different G $\alpha$  subunits, G $\alpha$ s possesses an aspartate at position 229, whereas all other G $\alpha$  subunits possess a serine at this position. As one of the mutants had altered this aspartate to a valine, suggesting its importance in  $\beta$  $\gamma$  binding, it was asked whether simply changing the aspartate to the more common serine residue would also serve to improve the  $\alpha$  $\beta$  $\gamma$  interaction, while preserving a "natural" Switch 2 structure. Figure 2 shows that the D229S allele confers a phenotype nearly as dramatic as that conferred by the original  
10 D229V mutation. It is also worthy of note that one of the other mutations, N254D, is analogous to D229S in that all G $\alpha$  subunits except that G $\alpha$ s possess an aspartate at this  
15 position.

In order to improve further on the  $\alpha$  $\beta$  $\gamma$  interaction, a double mutant E10K  
20 D229S was constructed using standard molecular biology techniques. In Figure 2 it can be seen that the two mutations act synergistically in that the double mutant possesses a phenotype drastically improved over that of either single mutation, to a level comparable with the endogenous yeast Gpal.

25 In order to avoid the inherent instability (~5% per cell division) of episomal plasmids in yeast, mutant G $\alpha$ s genes were integrated (E10K, D229S, S286P, and N254D, as well as the double mutant E10K D229S) at the gpal locus of CY1316, thereby preserving the promoter and terminator used in the episomal version. The resultant strains mimic the behavior of the analogous strains carrying mutant G $\alpha$ s on an  
30 episome.

In order to show that the mutant G $\alpha$ s subunits can not only associate tightly with yeast  $\beta$  $\gamma$ , but also dissociate from  $\beta$  $\gamma$  upon receptor stimulation, the human  
35 adenosine A2a receptor was introduced under the control of the PGK promoter into the various integrated G $\alpha$ s mutant backgrounds, along with an episomal fusl-lacZ reporter plasmid. As expected, the addition of the exogenous adenosine analog NECA results in

a ~10 fold increase in  $\beta$ -galactosidase units, demonstrating the ability of the mutant  $G\alpha s$  to interact productively with receptor, and to dissociate from  $\beta\gamma$  upon ligand addition.

The S286P mutation suggested a possible method for expanding the spectrum of  $G\alpha$  subunits that can interact with yeast  $G\beta\gamma$ . It was reasoned that because all mammalian  $G\alpha$  subunits possess a serine at this position and a proline is found in G $\alpha$ 1, the S286P substitution might be extrapolated to other mammalian  $G\alpha$  subunits to improve their apparent  $G\beta\gamma$  affinity. It has previously been shown that the substitution of the first 43 amino acids of yeast G $\alpha$ 1, corresponding to the amino terminal alpha-helical domain, for the corresponding segment of G $\alpha$ i2 results in a dramatic increase in the apparent affinity of G $\alpha$ i2 for the yeast  $G\beta\gamma$  subunit, as revealed by suppression of the constitutive  $G\beta\gamma$  pheromone signal ("41-i2"). This modification to G $\alpha$ 16 and G $\alpha$ 12 improved their ability to sequester yeast  $\beta\gamma$ , and this affinity is further strengthened by the substitution of the appropriate serine residue (corresponding to position 286 in  $G\alpha s$ ) with proline. In addition, the apparent  $\beta\gamma$  affinity of the triple chimera G $\alpha$ 1(1-43)-G $\alpha$ i2(36-242)-G $\alpha$ B(243-354) was also significantly improved by the appropriate S → P substitution.

20 Example 3. Construction of STE18/ mammalian  $G\gamma$  chimeric proteins

Chimeric proteins comprising STE18, the wild type  $G\gamma$  subunit of yeast, were also made. The wild type STE18 nucleotide sequence is available in the art. The carboxy terminal 34 amino acids of STE18 are as follows:

25 GYPVAGSNHFIEGLKNAQKNSQMSNSNSVCCTLML (SEQ ID  
NO: 29)

The wild type human  $G\gamma$  nucleotide sequences are available in the art. The carboxy terminal 24 amino acids of human  $G\gamma$  are as follows:

30 DPLLTPVPASENPFREKKFFCAIL (SEQ ID NO: 30)

The underlined residues shown are conserved among all of the mammalian  $G\gamma$  subunits

35 The STE18-mammalian  $G\gamma$  chimeras were constructed using standard molecular biology techniques. The following is a list of the  $G\gamma$  chimeras which were

generated, all are modified forms of yeast STE 18 which comprise a portion of a  $\text{G}\gamma$  subunit. The sequences shown below represent a carboxy terminal alignment of the sequences. The two dots before the first amino acid shown in each of the sequences in the following list indicate that the wild type STE18 amino terminal sequence is present 5 in all cases before the first amino acid shown. The residues of yeast STE18 origin are in normal type, while those of mammalian  $\text{G}\gamma$  origin appear in bold. The last amino acid shown represents the carboxy terminus of each chimera.

	Name of Chimera	Amino Acid Sequence	Mammalian $\gamma$ subunit
10		S $\gamma$ 0 ..GYPVAGSNHFIEGLKNAQKNSQMSNSNSV <b>CAIL</b>	G $\gamma$ 2
	E-/+ (SEQ ID NO:31)		
15	ID NO:32)	S $\gamma$ 4 ..GYPVAGSNHFIEGLK <b>F</b> FCCTL M	G $\gamma$ 2 E <sup>+</sup> (SEQ
	(SEQ ID NO:33)		
20		S $\gamma$ 6 ..GYPVAGSNHFIEGLKNP <b>FREKKFF</b> CCTL M	G $\gamma$ 2 E <sup>+</sup>
	(SEQ ID NO:34)		
25		S $\gamma$ 10 ..GYPVAGSNP <b>FREKKFF</b> CCTL M	G $\gamma$ 2 E <sup>+</sup>
	(SEQ ID NO:35)		
30		S $\gamma$ 12 ..GYPVAGSNHF <b>EKKFF</b> CCTL M	G $\gamma$ 2 E <sup>+</sup>
	(SEQ ID NO:36)		
35		S $\gamma$ 14 ..GYPVAGSNHF <b>EKKFF</b> CCTL M	G $\gamma$ 2 E <sup>+</sup>
	(SEQ ID NO:37)		
		S $\gamma$ 601 ..GYPVAGSNHFIEGLKNP <b>FKELKGG</b> CCTL M	G $\gamma$ 1 E <sup>+</sup>
	(SEQ ID NO:38)		
		S $\gamma$ 605 ..GYPVAGSNHFIEGLKNP <b>FRPQKV</b> CCTL M	G $\gamma$ 5 E <sup>+</sup>
30	(SEQ ID NO:39)		
		S $\gamma$ 607 ..GYPVAGSNHFIEGLKNP <b>FKDKKP</b> CCTL M	G $\gamma$ 7 E <sup>+</sup>
	(SEQ ID NO:40)		
		S $\gamma$ 610 ..GYPVAGSNHFIEGLKNP <b>FREPRSC</b> CCTL M	G $\gamma$ 10 E <sup>+</sup>
	(SEQ ID NO:41)		
35		S $\gamma$ 611 ..GYPVAGSNHFIEGLKNP <b>FKEKGSC</b> CCTL M	G $\gamma$ 11 E <sup>+</sup>
	(SEQ ID NO:42)		

S<sub>γ</sub>16 ..GYPVAGSNHFIEGLKNAQKNPFREKKFFCCTLM G<sub>γ</sub>2  
E<sup>+</sup> (SEQ ID NO:43)  
S<sub>γ</sub>18 ..GYPVAGSNHFIEGLKNPFREKKFFCTIL G<sub>γ</sub>4 E<sup>+</sup>  
5 (SEQ IDNO:44)

#### Example 4. Screens for modulators of G Protein activity

10 Screens for modulators of G protein, in this case G<sub>α</sub>, activity may be performed as shown in the following examples for illustration purposes, which are intended to be non-limiting.

15 Strains CY4874 and CY4877 are isogenic but for the presence of Q205L mutation in the cloned G<sub>αi2</sub> gene cloned into plasmid 1. Strains CY4901 and CY4904 each have a chromosomally integrated chimeric G<sub>α</sub> fusion comprising 41 amino acids of gpal at the N terminus of the human G<sub>αi2</sub> gene and are isogenic but for the presence of a constitutively activating mutation in the C5a receptor gene of CY4901. Strain CY5058 is a gpal mutant which carries only the yeast G<sub>βγ</sub> subunits and no G<sub>α</sub> subunit. This strain is a control strain to demonstrate specificity of action on the G<sub>α</sub> subunit.

20

#### Suppression of Activation by Mutation of G<sub>α</sub>

The Q205L mutation is a constitutively activated GTPase deficient mutant of the human G<sub>αi2</sub> gene. Antagonist compounds, chemicals or other substances which act on G<sub>αi2</sub> can be recognized by their action to reduce the level of activation and thus reduce the signal from the fus1-lacZ reporter gene on the second plasmid (Plasmid 2).

#### A. GTPase G<sub>αi2</sub> Mutants

test component = gpa41-G<sub>αi2</sub> (Q205L)

30 control component = gpa41-G<sub>αi2</sub>

As well as the CY4874 and CY4877 constructs detailed above, similar strains with fus1-His3 or fus2-CAN-1 growth readouts may also be used. The fus1-His3 strains are preferred for screening for agonists and the fus2-CAN1 strains are preferred for antagonist screens.

- 89 -

<u>Readout</u>	<u>test</u> <u>strain</u>	<u>effect of Gα<sub>i2</sub> antagonist</u>	<u>control</u> <u>strain</u>
fus1-HIS3	CY4868	inhibit growth of -HIS +AT (Aminotriazole)	CY4871
fus1-lacZ	CY4874	reduce β-gal activity	CY4877
fus2-CAN1	CY4892	induce growth on canavanine	CY4386

In each case an antagonist should cause the test strain to behave more like the control strain.

5

#### B. GTPase Gα<sub>s</sub> Mutants (Gα Specificity)

test component = Gα<sub>s</sub>(Q227L)

control component = Gα<sub>s</sub>

<u>Readout</u>	<u>test</u> <u>strain</u>	<u>effect of Gα<sub>i2</sub> antagonist</u>	<u>control</u> <u>strain</u>
fus1-HIS3	CY4880	none	CY4883
fus1-lacZ	CY4886	none	CY4889
fus2-CAN1	CY4895	none	CY4898

In each case a non-specific antagonist would cause the test strain to behave more like the control strain.

10

Additional media requirements: -TRP for Gα plasmid maintenance in fus1-HIS3 and fus2-CAN1 screens and -TRP -URA for Gα and fus1-lacZ plasmid maintenance in fus1-lacZ screen.

15

#### II. Suppression of Activation by Receptors

##### Constitutively Activated C5a Receptors

test component = C5aR\* (P184L activated C5a Receptor)

20

control component = C5aR

The C5aR\* mutation has a Leucine residue in place of the Proline residue of the wild-type at position 184 of the amino acid sequence.

- 90 -

<u>Readout</u>	<u>test strain</u>	<u>effect of Gα<sub>i2</sub> antagonist</u>	<u>control strain</u>
fus1-HIS3	CY4029	inhibit growth of -HIS +AT (Aminotriazole)	CY2246
fus1-lacZ	CY4901	reduce β-gal activity	CY4904
fus2-CAN1	CY4365	induce growth on canavanine	CY4362

In each case an antagonist should cause the test strain to behave more like the control strain.

- 5 Additional media requirements: -LEU for receptor plasmid maintenance in fus1-HIS3 and fus2-CAN1 screens and -LEU-URA for receptor and fus1-lacZ plasmid maintenance in fus1-lacZ screen, non-buffered yeast media (pH 5.5).

10 The data presented below show a Summary of various Gα Families  
Coupled in Yeast

			min[AT]i (mM)	Coupling
15	Gαi:	GPA1(41)-Gαi2	0.5	yes
		GPA1(41)-Gαi3	0.5	yes
		Gαi1	1.5	yes
		GPA1(41)-Gαi1	0.5	yes
20	Gαs:	various point mutants	0-0.5	yes
	Gα16:	GPA1(41)-Gα16(S-P)	1.0	yes
		GPA1(20)-Gα16	1.0	yes
		GPA1(20)-Gα16(S-P)	0.5	yes
25	Gαo:	GPA1(41)-Gαi2-Gαo(40)	1.0	yes
		GPA1(41)-Gαi2-Gαo(110)(S-P)	1.0	yes
		GPA(41) with STE4Dins	0.5	
Gα12/13:	GPA1(41)-Gα12	2.5-5.0	LIRMA	
	GPA1(41)-Gα12(S-P)	1.0		

5      **Example 5. The S $\gamma$ 6 mutant is effective in coupling to mammalian G protein  
coupled receptors**

The effect of the S $\gamma$ 6 mutant was tested on the coupling of mammalian GPCRs. Results are shown which compare the ability of the wild-type STE18 and S $\gamma$ 6 mutant in the context of four different G $\alpha$ s. Note that GPA41G $\alpha$ 12 is a chimera 10 containing the amino terminal 41 amino acids of yeast GPA1 in tandem with G $\alpha$ 12; GPA41G $\alpha$ 16(S27P) contains the amino terminal 41 amino acids of yeast GPA1 fused to G $\alpha$ 16 which has a Pro for Ser substitution at position 27; G $\alpha$ s(D229S) is the mammalian G $\alpha$ s with a Ser for Asp substitution at position 229.

Receptor	gpa41G $\alpha$ i2		gpa41G $\alpha$ 16(S27P)		G $\alpha$ s(D229S)		GPA1	
	STE18	S $\gamma$ 6	STE18	S $\gamma$ 6	STE18	S $\gamma$ 6	STE18	S $\gamma$ 6
FPR1 and Galanin	-	+++	-	-	-	-	-	-
Rat VIP	-	-	-	-	-	+++		
ML1b	-	-	-	-	-	-	++	
C5a	+++	+++	+/- (L)	++	-	-	-	
FPR1	+++	+++	+(L)	++	-	-	-	
IL-8 (rabbit)	+++	+++	LIRMA	++	-	-	++	
A2a	-	-	++(L)	++	+++	-	++	

15      (L) indicates a low level of ligand independent receptor activation (LIRMA) that does not obscure ligand dependent signal.

LIRMA indicates a high level of LIRMA that totally obscures any ligand dependent effect

+ indicates ligand dependent growth

20      - indicates no ligand dependent growth

The following abbreviations are used in the Table. VIP (vasoactive intestinal peptide); ML1b (melatonin receptor); C5a (complement cascade component); FPRL (formyl peptide related receptor); IL-8 (interleukin 8); A2a (Adenosine 2a receptor)

5

**Example 6. Development of four yeast strains for orphan receptor expression.**

This Example illustrates the development of four yeast strains, each expressing different chimeric G protein subunits, for use in drug screening assays. The use of four different types of G protein subunits in the yeast cells provides an

10 opportunity to achieve optimal G protein receptor coupling in at least one of the yeast strains. The genotypes of four exemplary yeast strains are illustrated in the table below:

Genotypes of Four exemplary yeast strains for Orphan Receptor Expression

CY10103	MAT $\alpha$ ste 18 $\gamma$ 6-3841 gpal (41)-G $\alpha$ i2 far1 $\Delta$ 1442 cyh2 tbt1-1 fus1-HIS3 can1 ste14::trp1::LTS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3 ade2 $\Delta$ 3447 ade8 $\Delta$ 3457
CY10132	MAT $\alpha$ ste 18 $\gamma$ 6-3841 gpal (41)-G $\alpha$ 16(S270P) far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3 ade2 $\Delta$ 3447 ade8 $\Delta$ 3457
CY10150	MAT $\alpha$ ste 18 $\gamma$ 6-3841 gpal (41)-G $\alpha$ s(D229S) far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3 ade2 $\Delta$ 3447 ade8 $\Delta$ 3457
CY10560	MAT $\alpha$ ste 18 $\gamma$ 6-3841 far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 sst2 $\Delta$ ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3 ade2 $\Delta$ 3447 ade8 $\Delta$ 3457

15 As discussed above, G $\alpha$  and G $\gamma$  chimeras are integrated at the gpal and ste18 loci respectively. Fus1-HIS3 is integrated at the fus1 locus and is phenotypically fus1 minus. A listing of phenotypes associated with the genotypes listed above is provided below.

MAT $\alpha$	mating type resulting in production of $\alpha$ -factor and responsiveness to $\alpha$ -factor
ste18 $\gamma$ 6	chimeric yeast G $\gamma$ /Human G $\gamma$ 2, enhances interaction with receptor
gpal(41)-G $\alpha$ i2	chimeric yeast G $\alpha$ /Human G $\alpha$ i2, enhances interaction with receptor
gpal(41)-G $\alpha$ 16(S270P)	chimeric yeast G $\alpha$ /mutant Human G $\alpha$ 16, enhances interaction with yeast $\beta$ y
gpalp-G $\alpha$ s(D229S)	mutant Human G $\alpha$ s, enhances interaction with yeast $\beta$ y
far1 $\Delta$ 1442	eliminates growth arrest response in pheromone pathway
cyh2	recessive resistance to cycloheximide

tbt1-1	poorly characterized enhanced transformation by electroporation
fus1-HIS3	pheromone responsive histidine prototrophy (aminotriazole resistance)
can1	recessive resistance to canavine
ste14::trp1::LYS2	eliminates carboxymethylation of isoprenylated proteins leading to reduced background through the pheromone response pathway
sst2Δ2	supersensitivity resulting from GAP activity on GPA1
ste3Δ1156	deletion of a-factor receptor gene
lys2	lysine auxotrophy and resistance to α-amino adipate
ura3	uracil auxotrophy, complementation by ligand plasmids
leu2	leucine auxotrophy, complementation by receptor plasmids
trp1	tryptophan auxotrophy, complementation by fus1-lacZ plasmid
his3	histidine auxotrophy, complementation by fus1-HIS3
ade2Δ3447	adenine auxotrophy, leads to generation of red pigment
ade8Δ3457	adenine auxotrophy, eliminates generation of red pigment in ade2 cells. complemented by receptor plasmid leading to colorimetric verification of its presence

#### Example 7: Construction of the α-Factor Leader-based Expression Vectors

- 5 A yeast vector for the expression of mammalian G protein-coupled receptors fused to a leader sequence of prepro-α-factor has been constructed as follows. A 0.38-kb fragment including a transcription terminator of yeast gene PH05 was amplified by PCR using a plasmid pTER as a template. The latter plasmid was constructed by subcloning of a Sau3A-PstI fragment of PH05 gene (GenBank accession number A07173) into the vector pUC19 digested with BamHI and PstI. PCR primers used were TER1, 5'-GGATCTAGAGGATCCTGGTACGTTCTC-3' (SEQ ID NO: 6), and TER2, GTCGCTAGCCAAGCTTGCATGCCCTGCAG-3' (SEQ ID NO: 7) (BRL, Life Technologies, Gaithersburg, MD). The primers provided XbaI and NheI restriction sites at the 5'-and 3'-terminus of the amplified fragment, respectively. The total of 30 cycles of PCR was performed; each cycle included denaturation at 94°C for 45 sec, annealing at 53°C for 1 min, and polymerization at 72°C for 1.5 min. The amplified fragment was digested with XbaI and NheI and subcloned into XbaI site of the plasmid Cadus 1289 (pLPX<sub>T</sub>) (LEU2 PGKp 2μ ori REP3 AmpR) in appropriate orientation.
- 10
- 15

- 94 -

This gave rise to the plasmid Cadus 4257 (pPPS) (LEU2 PGKp PHO5t 2mu-ori REP3 AmpR).

A 0.27-kb fragment of yeast gene MF $\alpha$ 1 encoding the entire prepro-region and a first spacer peptide KREAEA (SEQ ID NO: 86) of the  $\alpha$ -factor precursor was amplified by PCR using a plasmid pAC109 (Ostanin et al. 1994. J. Biol. Chem. 269:8971) as a template together with a pair of oligonucleotide primers, MF5, 5'-GCAGTCATGAGATTCTCAATTTACTGC-3', (SEQ ID NO: 8) and MF3, 5'-CAGCCCATGGCTTCAGCCTCTCTTATCC-3' (SEQ ID NO: 9). PCR was performed under the same conditions as those described above for the amplification of PHO5 terminator. As a result, the BspHI and NcoI restriction sites were generated at the 5'-and 3'-terminus of the amplified fragment, respectively. The fragment was treated with BspHI and NcoI and subcloned into an NcoI site of plasmid Cadus 4257 in appropriate orientation.

The resulting plasmid Cadus 4258 (pPMP15) (LEU2 PGKp Mf $\alpha$ 1 prepro PHO5t REP3 AmpR) carries a strong constitutive promoter of PGK gene followed by the  $\alpha$ -factor leader-encoding sequence, as well as the PHO5 transcription terminator. The unique restriction sites NcoI situated at the 3'-terminus of the leader sequence, as well as XbaI and BamHI localized at the 5'-terminus of the terminator can be used for subcloning of the G protein-coupled receptor-encoding sequences. The presence of a transcription terminator in the expression vector may allow avoidance of heterogeneity of the receptor specific mRNA and, therefore, to increase its stability.

In addition, a vector Cadus 4431 (pPMP15-HA) has been constructed which can be used for the expression of receptors tagged with hemagglutinin (HA) epitope at their C-termini. A synthetic adaptor composed of two oligonucleotides, HAtop, 5'-GATCCGCTTACCCATACGATGTTCCAGATTACGCTGCTTGA-3' (SEQ ID NO: 10)', and HAbot, 5'-GATCTCAAGCAGCGTAATCTGGAACATCGTA-3' (SEQ ID NO: 11)', was introduced into a BamHI site of the plasmid Cadus 4258 in the orientation that restores this site only on the side proximate to the leader-encoding sequence. Insertion of the receptor-encoding sequence lacking a stop codon into the NcoI and BamHI sites of the resulting vector encode a receptor bearing a C-terminal extension SAYPYDVPDYAA (SEQ ID NO: 12).

**Example 8. Development of a Functional Assay for the Human Nociceptin Receptor**

The prepro sequences of yeast  $\alpha$ -factor were inserted 3' to the PGK promoter and just 5' to the receptor cloning site of the standard yeast expression vector

pLPXt (a PGK promoter expression vector, using the NcoI and XbaI sites) to create the vector pMP15. This results in the synthesis of a chimeric receptor that matures to receptor devoid of  $\alpha$ -factor sequences in the Golgi.

5     This vector was used to develop a yeast-based human nociceptin receptor assay. The prepro- $\alpha$ -factor-nociceptin receptor expression plasmid was introduced into several isogenic strains differing in expression of chimeric STE18-human Gy proteins, as described herein.

10    The nociceptin receptor expressed as a fusion to the  $\alpha$ -factor leader exhibited a cell surface staining pattern which was similar to that observed for the Ste2 receptor. Consistent with this, the fusion receptor was shown to undergo both outer chain glycosylation and processing by Kex2 protease that occur in Golgi. Thus, the leader sequence of the  $\alpha$ -factor precursor appears to direct efficient transport of the receptor to the cell surface.

15    Receptor activity was also tested in an additional Cadus yeast strain, CY2120. CY2120 contains a deletion of the endogenous yeast G $\alpha$  protein GPA1, but carries the intact yeast Gy. Similar to other Cadus yeast strains, the endogenous HIS3 gene is defective and FUS1-HIS3 is integrated into the genome. This strain contains a novel mutation in the SST2 gene. SST2 encodes a protein that increases the rate at which GPA1 undergoes GTP hydrolysis. Thus in the absence of SST2, GPA1 stays in 20 the GTP-bound state longer. Functionally, this means that GPA1 is active for a longer period of time and is therefore capable of transmitting the signal from a seven transmembrane receptor for a longer period of time (if the particular seven transmembrane receptor is capable of interacting with GPA1).

25    This discovery was further corroborated by the results of a functional growth assay. The strain CY2120 (MAT $\alpha$  sst2\*2far1\*1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3\*1156 lys2 ura3 leu2 trp1 his3) coexpressing the  $\alpha$ -factor leader-nociceptin receptor fusion protein and the Gpa1G $\alpha$  subunit have been shown to exhibit nociceptin-induced growth as well as approximately 10-fold activation of the  $\beta$ -galactosidase expression in response to nociceptin. No ligand-induced growth was observed for the same strain 30 expressing the nonfusion receptor.

#### **Example 9. Improvement of a Growth Assay for the Human Melanocortin 4 Receptor.**

35    The human melanocortin 4 receptor (MC4R) was shown to be functional in yeast when expressed as a nonfusion protein. It couples to the mating pheromone response pathway through G $\alpha$ <sub>S</sub> D229S subunit, as demonstrated by growth activation of

strain CY9438 (MAT $\alpha$  ste18y6-3843 far1\*1442 tbt1-1fus1-HIS3 can1 ste14::trp1::LYS2 ste3\*1156 lys2 ura3 leu2 trp1 his3) in response to six different MC4R-specific agonists. However, the sensitivity of the growth assay is considerably improved when this receptor is expressed as a fusion to the  $\alpha$ -factor leader, presumably, due to the elevated level of the receptor at the cell surface.

The human Melanocortin 4 receptor cDNA was cloned into three yeast expression vectors under control of the PGK promoter. The full-length receptor cDNA expressed from the plasmid pLPXt was used in all biological-response assays. The plasmid pMP15, which appends the prepro sequences of  $\alpha$ -factor to the N-terminus of the receptor was also used in a subset of biological-response experiments. Finally, the plasmid pLPXt-FLU was used to append a hemagglutinin epitope to the C-terminus of the receptor. This final construct was utilized to verify MC4 receptor expression by Western blot analysis.

The melanocortin receptors have been demonstrated to interact with multiple ligands arising from the processing of pro-opiomelanocortin. All of these natural ligands are relatively small peptides.  $\alpha$ -MSH is a 13 amino acid peptide acylated at the amino-terminus and amidated at the carboxyl-terminus. A synthetic analog of  $\alpha$ -MSH with enhanced resistance to proteolysis is [Nle4, D-Phe7]- $\alpha$ -MSH (NDP-MSH). The lower potency agonists include  $\beta$ -MSH, which is 22 amino acids in length, and the family of  $\gamma$ -MSH peptides which vary in length from 11-27 amino acids (11 amino acid  $\gamma$ 1, 12 amino acid  $\gamma$ 2, and 27 amino acid  $\gamma$ 3).

The MC4 receptor is well characterized as signaling through G $\alpha$ s. Therefore, upon confirmation of receptor expression by Western blot, the full-length MC4 receptor was introduced along with several G $\alpha$ s expression plasmids, into a series of isogenic strains varying only in G $\gamma$  chimera composition. Similarly, MC4 receptor expressed with the prepro- $\alpha$ -factor leader was introduced into one of the G $\gamma$  chimera strains with the series of G $\alpha$ s expression plasmids.

Ligands were applied to cell monolayers, and receptor activation was examined by the FUS1-HIS3 growth assay. Upon ligand addition, the unmodified receptor proved to be most active in CY9800. Only two ligands however, were capable of receptor stimulation,  $\alpha$ -MSH and NDP-MSH. In contrast, when  $\alpha$ -factor prepro leader sequences were appended to the N-terminus of the receptor all ligands activated the receptor in the sole strain assayed, CY9438. Quantitation of ligand-dependent activation of the yeast mating pathway was measured through induction of FUS1- $\beta$ -galactosidase in a non-optimized microtitre format. In this assay, maximal stimulation was roughly 5-fold with an EC<sub>50</sub> of approximately 12 nM.

**Example 10. Yeast Expression of the Rat Metabotropic Glutamate Receptor 2(mGluR2).**

5           Metabotropic glutamate receptors, in contrast to other seven transmembrane receptors, possess the unusually long first extracellular domains. Specifically, the N-terminal domain of rat mGluR2 consists of 567 amino acid residues.

10          The rat mGlu2 receptor cDNA was sequenced and found to be identical to that in the published literature (GenBank accession number D16817). Thereafter, the cDNA was subcloned into eight yeast expression vectors.

15          The plasmid pLPXt was modified for expression of the mGlu2 receptor from the PGK promoter. Transcription of receptors in this vector is normally terminated approximately 400 nucleotides 3' from the end of the cDNA insert. Increasing evidence however, suggested that it might be possible to increase the steady state level of transcript by including a transcriptional terminator adjacent to the carboxyl-terminus of the cDNA. Thus pLPXt was modified by the addition of a transcriptional terminator from the PHO5 gene. This construct is referred to as mGlu2. Two constructs were also prepared using the previously described prepro- $\alpha$ -factor leader. In the construct designated pp $\alpha$ F-mGluR2, the leader sequences were inserted adjacent to the N-terminus of the receptor. In contrast, the postulated signal sequence of the mGlu2 receptor was deleted, and replaced by the prepro- $\alpha$ -factor leader in the construct pp $\alpha$ F-mGluR2 $\Delta$ . Finally, the construct invmGluR2 directs synthesis of the mGlu2 receptor with the leader of the invertase protein appended to the N-terminus of the receptor. Like the  $\alpha$ -factor leader, this leader has also been demonstrated to promote entrance of heterologous proteins into the secretory pathway of yeast. These four constructs were used to investigate the biological activity of the mGlu2 receptor in yeast. In order to evaluate expression by Western blot, the mGlu2 receptor was also cloned into a set of vectors identical to those just described, but with the hemagglutinin epitope attached to the carboxyl-terminus of the protein.

30          Plasmids expressing hemmagglutinin-tagged receptors were transfected into yeast, and receptor levels were determined by Western blot analysis of total membrane protein preparations. In the absence of yeast leader sequences no receptor was detected, even with prolonged exposure of the nitrocellulose membrane. In contrast, when expressed with the prepro- $\alpha$  factor or invertase leader, dark doublet bands of high molecular weight were detected. These results indicated that receptor expression could be dependent upon the presence of yeast sequences.

There was however, an alternative interpretation. Large, highly glycosylated proteins transfer to nitrocellulose with low efficiency. Therefore, it was a formal possibility that receptor lacking a yeast leader was hyper-glycosylated, rendering it resistant to detection by Western blot. To address this possibility, protein samples 5 were treated with Endoglycosidase H (Endo H), which cleaves the high mannose structures on N-linked oligosaccharides. Removal of oligosaccharides did not improve visualization of the leader receptor, suggesting that in the vector, the mGlu2R was not synthesized. Furthermore, the Endo H results indicated the presence of oligosaccharides on receptor made from constructs with leaders, once again indicating transport into the 10 secretory pathway. Together, these data underscore the important role yeast leader sequences play in mGlu2 receptor expression.

Several agonists are available for research on the mGlu2 receptor. The agonists used in these studies included L-CCG-I, 1S,3R-ACPD, and L-glutamic acid, with EC<sub>50</sub> values of 0.75 μM, 7.7 μM and 11.8 μM, respectively. Stimulation of the 15 mGlu2 receptor has been demonstrated to promote inhibition of adenylyl cyclase, thus initial attempts to couple the receptor focused primarily upon the use of proteins from the G<sub>αi</sub> family.

CY9437 expresses a unique Ste18-G<sub>γ</sub>2 chimera. To date, most receptors 20 that have been developed into successful assays have shown the ligand independent receptor activation: LIRMA, phenotype in this strain when presented with an appropriate G<sub>α</sub> subunit. Therefore in the absence of receptor coupling, CY9437 can be loosely used as a diagnostic indicator of the capacity of a receptor to couple in yeast. When compared with other mGlu2 receptor constructs, only ppaF-mGluR2Δ induced LIRMA. However, it is possible that rather than LIRMA, this is actually ligand-dependent 25 activation arising from glutamate released into the growth media by the yeast. Glutamate oxidase, which converts glutamate to α-ketoglutarate, was used in an attempt to diminish glutamate levels. Nonetheless, additional glutamate metabolizing enzymes (which may be more potent under yeast growth conditions) are available including L-glutamate decarboxylase and L-glutamate dehydrogenase.

30

#### **Example 11. Construction of G<sub>α1</sub> carboxy terminal chimera.**

The C-terminus of G protein α subunits have been shown to be crucial in determining receptor-G protein specificity (see e.g., Conklin *et al.*, (1995) *Nature* 363, 274-276; Liu *et al.*, (1995) *Proc. Natl. Acad. Sci.* 92, 111642-1116460; Hamm *et al.*, 35 (1998) *J. Biol. Chem.* 273:669-672; Conklin *et al.*, (1996) *Mol. Pharmacol.* 50:885-890). To test the ability for specific coupling through heterologous receptors expressed in yeast,

- 99 -

chimeric proteins of G<sub>p</sub>α1 were constructed. The last C-terminal 5 or 6 amino acids of G<sub>p</sub>α1 were substituted with the last 5 or 6 amino acids of mammalian G<sub>α</sub> subunits. To generate chimeric constructs, Cadus 1179, encoding the wild type G<sub>p</sub>α1, was used as a template for PCR amplification. Specific primers pairs for amplification are as follows:

5 :

GPA1-G<sub>α</sub>s(5) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC3' (SEQ ID NO: 87)  
Primer 2: 5' ACG TCT CGA GGT CGA CTC ATA GAA GCT CAT ATT  
GTT TAA GG3' (SEQ ID NO: 88)

10

GPA1-G<sub>α</sub>i2(5) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC3' (SEQ ID NO: 87)  
Primer 3: 5' ACG TCT CGA GAT CGA CTC AAA ATA GAC CAC  
AGT CTT TAA GGT TTT GCT GGA TG3' (SEQ ID NO: 89)

15

GPA1-G<sub>α</sub>Z(5) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC3' (SEQ ID NO: 87)  
Primer 4: 5' ACG TCT CGA GGT CGA CTC AAC AAA GAC CAA  
TAT ATT TAA GGT TTT GCT GGA TG3' (SEQ ID NO: 90)

20

GPA1-G<sub>α</sub>q(5) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC3' (SEQ ID NO: 87)  
Primer 5: 5' ACG TCT CGA GGT CGA CTC ATA CAA GAT TAT ATT  
CTT TAA GGT TTT GCT GGA TG3' (SEQ ID NO: 91)

25

GPA1-G<sub>α</sub>12(5) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC3' (SEQ ID NO: 87)  
Primer 6: 5' ACG TCT CGA GGT CGA CTC ATT GAA GCA TAA  
TGT CTT TAA GGT TTT GCT GGA TG3' (SEQ ID NO: 92)

30

GPA1-G<sub>α</sub>13(5) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC3' (SEQ ID NO: 87)  
Primer 7: 5' ACG TCT CGA GGT CGA CTC ATT GAA GCA TAA  
GTT GTT TAA GGT TTT GCT GGA TG3' (SEQ ID NO: 93)

35

GPA1-G<sub>α</sub>16(6) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC3' (SEQ ID NO: 87)  
Primer 8: 5' ACG TCT CGA GGT CGA CTC AAA GAA GAT TAA  
TTT CAT CAA GGT TTT GCT GGA TGA TT3' (SEQ ID NO: 94)

GPA1-G<sub>α</sub>s(6) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC3' (SEQ ID NO: 87)

40

Primer 9: 5' ACG TCT CGA GGT CGA CTC ATA GAA GCT CAT ATT  
GTC TAA GGT TTT GCT GGA TG3' (SEQ ID NO: 95)

- 100 -

GPA1-G $\alpha$ (5) Primer 1: 5' TTT GAA GGG CCG TAT AAA GAC 3' (SEQ ID NO: 87)

Primer 10: 5' ACG TCT CGA GAT CGA CTC AAT ATA GAC CAC  
ATC CTT TAA GGT TTT GCT GG 3' (SEQ ID NO: 96)

Following amplification, the PCR products, which were approximately 5 550bp in size, were column purified using the Qiagen PCR purification kit. The purified amplification product was digested with BstBI and Xhol for cloning into Cadus 1179. Cadus 1179 was digested with BstBI and Xhol, treated with shrimp alkaline phosphatase and purified by gel electrophoresis. The PCR product and linearized vector were ligated and electroporated into competent bacteria. Colonies grown on ampicillin 10 plates were selected and amplified. The DNA was isolated from the bacterial cells and sequenced to provide the following C-terminal chimera:

15

**Alignment of GPA1 and mammalian G $\alpha$  subunits:**

G $\alpha$ 1: .....AVTDLIIQQNLKKIGII (SEQ ID NO: 97)  
ai2: .....AVTDVIIKNNLKDCGLF (SEQ ID NO: 98) (ai1 has the same last 5aa)

20 aq: .....AVKD TILQLNLKEYNLV (SEQ ID NO: 99)(all has the same last 5aa)  
  
a12: .....AVKD TILQENLKDIMLO (SEQ ID NO: 100)  
as: .....DCRD II QRMHLROYELL (SEQ ID NO: 101)  
a13: .....DV KDTILHDNLKQLMLO (SEQ ID NO: 102)  
25 az: .....AVTDVIIQNNLKKYIGLC (SEQ ID NO: 103)  
a16: .....DVRDSVLARYLDEINLL (SEQ ID NO: 104)  
as: .....DCRD II QRMHLROYELL (SEQ ID NO: 105)  
ao: .....AVTDIIIANNLRGCGLY (SEQ ID NO: 106)

30 The C-terminal amino acids of *S. cerevisiae* GPA1 and mammalian G $\alpha$  subunits are shown above. For each mammalian G $\alpha$  subunit shown, the amino acids which were exchanged with amino acids at the C-terminus of GPA1 are shown in bold and underlined.

35 The last 5 amino acids of GPA1 were substituted with the last 5 amino acids from the corresponding mammalian G $\alpha$  subunit to create the following GPA1-G $\alpha$ (5) chimeras:

GPA1-G $\alpha$ i2(5) (SEQ ID NO: 107)

GPA1-G $\alpha$ q(5), (SEQ ID NO: 108)

- GPA1-G $\alpha$ 12(5),(SEQ ID NO: 109)  
 GPA1-G $\alpha$ s(5),(SEQ ID NO: 110)  
 GPA1-G $\alpha$ 13(5),(SEQ ID NO: 111)  
 GPA1-G $\alpha$ z(5),(SEQ ID NO: 112)  
 5 GPA1-G $\alpha$ o(5) (SEQ ID NO: 113)

The last 6 amino acids of GPA1 were substituted with the last 6 amino acids from the corresponding mammalian G $\alpha$  subunit to create the following GPA1-G $\alpha$ (6) chimeras:

- 10 GPA1-G $\alpha$ 16(6)(SEQ ID NO: 114)  
 GPA1-G $\alpha$ s(6)(SEQ ID NO: 115)

A summary of chimeric G proteins is presented in Table 1. Sandwich chimeric G proteins are also shown in Table 1 and will be discussed in more detail in Example 12.

15

20

**Table 1      Gpa1 -G $\alpha$ (5), Gpa1-G $\alpha$  (6) and sandwich chimeras**

G protein	Cadus Plasmid Number
GPA1-G $\alpha$ i2(5)	Cp4921, 4920
GPA1-G $\alpha$ s(5)	Cp5007, 5008
GPA1-G $\alpha$ 12(5)	Cp5040, 5041
GPA1-G $\alpha$ q(5)	Cp5042, 5043
GPA1-G $\alpha$ z(5)	Cp5169, 5170
GPA1-G $\alpha$ 13(5)	Cp5407, 5408
GPA1-G $\alpha$ 16(6)	Cp5531, 5532
GPA1-G $\alpha$ s(6)	Cp5905
G $\alpha$ q(1-11)-GPA1(6-467)-G $\alpha$ q(355-359)	Cp5902
G $\alpha$ q(1-21)-GPA1(24-467)-G $\alpha$ q(355-359)	Cp6079
GPA1-G $\alpha$ o(5)	Cp6193

**Construction of Yeast Strains with Integrated G-proteins:**

25 GPA1 chimerae were integrated into the yeast genome such that the native GPA1 open reading frame was replaced with the chimeric open reading frame.

GPA1-G $\alpha$ i2(5), GPA1-G $\alpha$ s(5), GPA1-G $\alpha$ 12(5), and GPA1-G $\alpha$ q(5) were integrated in two steps (Boeke et al., Meth. Enzymol. 154: 164-175) into CY1316 (MAT $\alpha$  gpa1 $\Delta$ 1163 far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3) and CY2120 (MAT $\alpha$  sst2 $\Delta$ 2 gpa1 $\Delta$ 1163 far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3). GPA1-G $\alpha$ z(5), GPA1-G $\alpha$ 13(5), GPA1-G $\alpha$ 16(6), GPA1-G $\alpha$ s(6), G $\alpha$ q(1-11)-GPA1(6-467)-G $\alpha$ q(355-359), and G $\alpha$ q(1-21)-GPA1(24-467)-G $\alpha$ q(355-359) were integrated in one step (Erdeniz et al., Genome Res. 7:1174-83; Rothstein, Meth. Enzymol. 194: 281-301) into CY14014 (MAT $\alpha$  sst2 $\Delta$ 2 gpa1::klURA3 far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3) and CY14066 (MAT $\alpha$  ste18g6-3841 sst2 $\Delta$ 2 gpa1::klURA3 far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3).

Table 2 Shows the nomenclature of mammalian GPA1-G $\alpha$ (5), GPA1-G $\alpha$ (6) and sandwich chimeras integrated into yeast

G Protein	Cadus Yeast Number in STE18 sst2 $\Delta$	Cadus Yeast Number STE18-Sy6 sst2 $\Delta$
GPA1-G $\alpha$ i2(5)	CY12946, 12947	CY13393, 13394
GPA1-G $\alpha$ s(5)	CY12952, 12953	CY13399, 13400
GPA1-G $\alpha$ 12(5)	CY12948, 12949	CY13395, 13396
GPA1-G $\alpha$ q(5)	CY12950, 12951	CY13397, 13398
GPA1-G $\alpha$ z(5)	CY15347, 15348	CY15070, 15071
GPA1-G $\alpha$ 13(5)	CY15349, 15350	CY15074, 15075
GPA1-G $\alpha$ 16(5)	CY15351, 15352	CY15072, 15073
GPA1-G $\alpha$ s(6)	CY17018, 17019	CY17008, 17009
G $\alpha$ q(1-11)-GPA1(6-467)-G $\alpha$ q(355-359)	CY17012, 17013	CY17002, 17003
G $\alpha$ q(1-21)-GPA1(24-467)-G $\alpha$ q(355-359)	CY17020, 17021	CY17010, 17011

LacZ Assay of Melatonin 1a and FPRL1 Receptors:

CY1141 (MAT $\alpha$  far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 gpa1(41)-G $\alpha$ i2 lys2 ura3 leu2 trp1 his3), CY10981 (MAT $\alpha$  GPA+3907 sst2 $\Delta$ 2 far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3), and CY12946 (MAT $\alpha$  sst2 $\Delta$ 2 GPA1-G $\alpha$ i2(5) far1 $\Delta$ 1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3 $\Delta$ 1156 lys2 ura3 leu2 trp1 his3) were transformed with the following plasmids to create strains CY15502, CY15490, CY15494, CY6539, CY15438, CY15440, CY6545, CY15437 and CY15439:

CADUS 1584 + CADUS 2311 LacZ<sup>+</sup> FPRL1 Receptor<sup>+</sup>  
CADUS 1584 + CADUS 2695 LacZ<sup>+</sup> Melatonin 1a Receptor<sup>+</sup>  
CADUS 1584 + CADUS 1289 LacZ<sup>+</sup> Receptor<sup>+</sup>

5

A pool of transformants was picked from each transformation and grown overnight in media lacking tryptophan and leucine, at pH6.8 with 25mM PIPES. The optical density at 600 nm of a 1/10 dilution of the overnight cultures was determined and the cultures were diluted in fresh media to final OD<sub>600</sub> of 0.2. Strains with the  
10 melatonin receptor were then grown for an additional 1.5 hours and diluted again to an OD<sub>600</sub> of 0.2. The LacZ enzyme activity was determined at increasing concentrations of ligand. The LacZ enzyme assay was performed in a 96 well plate and each reaction was performed in triplicate in a total volume of 100 µl. For each reaction, 90 µl of culture and 10 µl of ligand were used. The final concentration of DMSO in each well was kept  
15 constant at 1% for the FPRL1 receptor and at 5% for the Melatonin 1a receptor.

LacZ activity, in the presence and absence of the FPRL1 receptor, was measured at the following concentrations of FPRL1: 0nM, 1.6nM, 8nM, 40nM, 200nM, 1µM, and 25µM. Following the addition of the ligand, the 96 well plates were incubated at 30C for 4 hours. 20µl of 0.5mM fluorescein di-β-D-galactopyranoside (FDG) was  
20 added as a substrate of the reaction. The plates were incubated at 37C for 90 mins. Following incubation, the reaction was stopped by the addition of 20ml 1M Na<sub>2</sub>CO<sub>3</sub> to each well. Plates were read using a fluorometer at an excitation wavelength of 485nm and an emission wavelength of 535nm.

LacZ activity, in the presence and absence of the melatonin receptor, was measured at the following concentration of melatonin: 0pM, 1pM, 10pM, 100pM, 1nM, 10nM, 100nM, 1µM, 10µM, 100µM, 1mM and 4.3mM. Results show that the G<sub>α1</sub>, α<sub>i2</sub> subunit and the C-terminal α<sub>i2</sub> (5) protein chimera couple and stimulate the FPRL1 receptor. With the melatonin receptor, stimulation of the receptor was observed with the chimeric G protein, as with wild type G<sub>α1</sub>. Coupling was also observed with the G<sub>α1</sub>-  
30 41α<sub>i2</sub> chimeric protein, although the stimulation of the receptor with this chimeric protein was less efficient. These results demonstrate that chimeric G proteins are able to couple and stimulate some heterologous receptors.

Stimulation through various other heterologous receptors was investigated using the chimeric G proteins. A summary of the results is presented in  
35 Table 3. Data from Table 3 shows that stimulation of heterologous receptors is observed with several chimeric G proteins in which G<sub>α1</sub> alone failed to stimulate.

Table 3 Coupling of mammalian receptors to G<sub>p</sub>a1-(c5) chimerae

Receptor		GPA1	GPA1-(c5 or 6) chimera			
			αi2(5)	αq(5)	α12(5)	αs(5)
ML1aR	Melatonin receptor 1a	+++	+++	-	+	ND
ML1bR	Melatonin receptor 1b	+++	+++	-	-	ND
NocR	Nociceptin receptor	+	+	-	-	ND
SSTR2	Somatostatin receptor 2	+	++	-	-	ND
FPR1 (1938)	Formyl peptide receptor 1	-	++	-	-	-
GalR1	Galanin receptor 1	-	++	-	-	-
rIL8R	Rabbit CXCR1	+++	+++	+	+	ND
C5aR	C5a receptor	+++	++	+	+	ND
A2aR	Adenosine receptor 2a	+++	++	-	-	+
rVIPR	Rat vasoactive intestinal receptor	-	-	-	-	-/+
MC4R	Melanocortin receptor 4	++	+	-	-	+
rNTR	Rat Neurotensin receptor	++	++	+++	++	ND
FPRL1	Formyl peptide receptor 2	++	++	-	-	-

Stimulation of heterologous receptors was also observed using G protein chimeras, or sandwich chimeric G proteins, in the Fus1 p-HIS3 assay, the results of which are presented in Table 4. The data demonstrates that many chimeric G proteins were able to stimulate heterologous receptors that wild type G<sub>p</sub>a1 was not able to stimulate.

Receptor	GPA1 and GPA1 chimeras	
	functional	nonfunctional
CXCR2 (mutant)	GPA1 GPA1-Gαi2(5) GPA1-Gα16(6)	
CXCR4	GPA1-Gαi2(5)	GPA1
Adenosine 2b receptor	GPA1-Gαs(5)	GPA1
Bombesin receptor subtype 3	GPA1-Gαz(5) Gαq(1-11)-GPA1(6-467)-Gαq(355-359) (FUS1p-HIS3)	GPA1 GPA1-Gαq(5)
Bradykinin receptor 2	GPA1-Gαq(5) Gαq(1-11)-GPA1(6-467)-Gαq(355-359)	GPA1

Coupling to heterologous orphan receptors using C-terminal chimeric G proteins or sandwich chimeras was tested using the lacZ assay. The results of the assay are shown in Table 5. Data demonstrates that stimulation via orphan receptors was observed using the chimeric G proteins (referred to as "tail" in Table 5). Stimulation was also observed with at the Gαq(1-11)-GPA1(6-467)-Gαq(355-359) sandwich chimera.

Table 5 Table showing orphan receptors and C-terminal chimeric G proteins or sandwich chimeras

Receptor	GPA1	i2 tail	12 tail	q tail	s tail	16 tail	13 tail	Q sand-wich (1-11)	S(6) tail	Q sand-wich (1-21)	z tail
Orphan 1	X	-	-	-	X*	-	-	-	-	-	-
Orphan 2	X	X	-	-	-	ND	-	ND	ND	ND	-
Orphan 3	X	X	X	-	-	ND	ND	ND	ND	ND	ND
Orphan 4	X	-	X*	-	-	ND	ND	ND	ND	ND	ND
Orphan 5	X	X	-	-	X	ND	-	ND	ND	ND	X
Orphan 6	-	-	-	X	X*	-	-	ND	ND	ND	-
Orphan 7	X	X	-	X	-	ND	ND	ND	ND	ND	ND
Orphan 8	X	-	-	-	-	-	-	ND	ND	ND	-
Orphan 9	-	X*	X*	-	-	-	X*	-	-	-	X*
Orphan 10	-	X*	X*	X*	-	ND	X*	ND	ND	ND	X*

	X	-	-	weak	weak	ND	-	X	-	-	-
Orphan 11											
Orphan 12	X	weak	weak	X*	-	ND	ND	ND	ND	ND	ND
Orphan 13	X	-	X	-	-	ND	-	ND	ND	ND	-
Orphan 14	X	X	-	weak	-	ND	X	ND	ND	ND	X
Orphan 15	-	-	X	-	-	-	-	weak	X	ND	-
Orphan 16	-	-	-	X*	-	-	-	X	ND	ND	-
Orphan 17	X	X	X	X	X	X	-	ND	ND	ND	X

X works

X\* works better in tails. ND: not done

#### Example 12. Construction of "sandwich" chimera G proteins.

In addition to the C-terminus of G protein  $\alpha$  subunits being crucial in determining receptor-G protein specificity, the N-terminus of G protein  $\alpha$  subunit also appears to be involved in receptor contact (see e.g., Hamm *et al.* (1998) J. Biol. Chem. 273:669-672). Sandwich chimeric proteins were constructed to investigate the effect on receptor specificity by substituting both the C-terminal and N-terminal regions of the G $\alpha$ 1 with protein G  $\alpha$  subunits.

#### Construction of G $\alpha$ q(1-11)-GPA1(6-467)-G $\alpha$ q(355-359)

The G $\alpha$ q(1-11)-GPA1(6-467)-G $\alpha$ q(355-359) sandwich chimera was constructed in which the first 11 amino acids of G $\alpha$ q and residues 6-381 of GPA1 was constructed using Cp1179 as a template. The following primers were used for PCR amplification:

Primer11:ACGTGGTCTCCATGACTTTGGAATCTATTATGGCTTGTGTCTTAG  
TACGCAAACAATAGGAGACG (SEQ ID NO: 116)

Primer12: GTATCTTGAACCACCTAGAG (SEQ ID NO: 117)

The PCR product containing the first 11 amino acids of G $\alpha$ q and residues 6-381 of GPA1, was purified and digested with BstB1 and Bsal. The digested product was ligated into Cp5042 (GPA1p-GPA1-G $\alpha$ q(5) CEN6 ARS4 AmpR TRP1). The ligation mixture was electroporated into competent bacteria. Plasmid DNA was prepared from ampicillin resistant bacterial colonies and the sequence of the G protein was verified by DNA sequence analysis. One clone which contained the 11 amino acids of G $\alpha$ q was identified and referred to as Cp5902: GPA1p-G $\alpha$ q(1-11)-GPA1(6-467)-G $\alpha$ q(355-359) CEN6 ARS4 AmpR TRP1, (SEQ ID NO: 118).

Construction of G $\alpha$ q(1-21)-GPA1(24-467)-G $\alpha$ q(355-359)

The G $\alpha$ q(1-21)-GPA1(24-467)-G $\alpha$ q(355-359) sandwich chimera was constructed in which residues 13-21 of G $\alpha$ q and residues 23-381 of GPA1 was constructed using Cp5902 as a template. PCR amplification was performed using two sets of primers. The first set of primers had the following sequence:

5 Primer13: GTCTAAAATGAAGAGGATAGTAG (SEQ ID NO: 119)

Primer14: GATCCGTCTCACTTCAGAAAGACAACAAAGCCATAATAG (SEQ ID NO: 120)

10 The first set of primers was used to generate a PCR product which contained the sequences of the GPA1 promoter downstream from the EcoRI site and the first 12 amino acids of G $\alpha$ q. The PCR product was purified and digested with EcoRI and BsmBI.

A second set of primers was used to amplify residues 13-21 of G $\alpha$ q and residues 23-381 of GPA1 using Cp5902 as a template and the following primers sequences:

15 Primer15: GATCCGTCTCTGAAGAAGCTAAGGAGGCTAGAAGAATTAAATGATG  
TCATCGAGCAATCGTTGC (SEQ ID NO: 121)

Primer12: GTATCTTGAACCACTTAGAG (SEQ ID NO: 122)

The resulting PCR product containing residues 13-21 of G $\alpha$ q and residues 23-381 of GPA1, was purified and digested with BsmBI and BstBI.

20 The above two digested PCR products were then ligated into Cp5042 (GPA1p-GPA1-G $\alpha$ q(5) CEN6 ARS4 AmpR TRP1), which was digested with EcoRI and BstBI. The ligation mixture was electroporated into competent bacteria. Plasmid DNA was prepared from ampicillin resistant bacteria and the sequence of the G protein was verified by DNA sequence analysis. One clone which contained the 21 amino acids 25 of G $\alpha$ q was identified and referred to as Cp6079: GPA1p-G $\alpha$ q(1-21)-GPA1(24-467)-G $\alpha$ q(355-359) CEN6 ARS4 AmpR TRP1 (SEQ ID NO: 123). A summary of the GPA1-G $\alpha$ (5), GPA1-G $\alpha$ (6) and Sandwich Chimeras constructs is provided in Table 1, in Example 11. Sandwich chimera G proteins integrated into yeast cells are shown in Table 2 in Example 11. Tables 4 and 5 in Example 11, show that sandwich chimeras are 30 able to couple to and stimulate heterologous receptors, determined using the Fus1-pHIS3 assay.

**GPA1-G $\alpha$ q sandwich improves functional activity of a bradykinin-responsive receptor:**

35 Construction of yeast strains with integrated G-proteins is described in Example 11. Sandwich chimera were tested for stimulation via heterologous receptors

by transforming CY2120 with human bradykinin receptor 2 under the control of the PGK promoter, FUS1p-lacZ reporter, and one of the following plasmids: Cp1179 (GPA1), Cp5042 (GPA1-G $\alpha$ q(5)), or Cp5902 (G $\alpha$ q(1-11)-GPA1(6-467)-G $\alpha$ q(355-359)). Transformants were grown to mid-exponential phase in medium lacking leucine, uracil, and tryptophan, at pH 6.8. 100  $\mu$ l aliquots of cultures were incubated in 96-well plates in the presence of increasing concentrations of bradykinin at 30°C for 4 hours. 20  $\mu$ l of 6X Z-buffer (0.6 M sodium phosphate, pH 7.0, 60 mM KCl, 6 mM MgCl<sub>2</sub>, 1.6% (v/v)  $\beta$ -mercaptoethanol) containing 10 mM chlorophenolred- $\beta$ -D-galactoside (CPRG) and 2.5% Triton X-100 was added to each aliquot followed by incubation at room 5 temperature for 60 min. The enzymatic reactions were stopped by adding 60  $\mu$ l of 1M sodium carbonate and A<sub>575</sub> values were determined using a Biomek Plate Reader (Beckman). Upon ligand treatment  $\beta$ -galactosidase activities were not elevated for transformants with Cp1179.  $\beta$ -galactosidase activities were increased 26 fold for transformants with G $\alpha$ q(1-11)-GPA1(6-467)-G $\alpha$ q(355-359); 13 fold for transformants 10 with GPA1-G $\alpha$ q(5).

15

#### Example 13. Detection using Fus1p-GFP

To provide an rapid, efficient method of detection, GFP was used as a reporter gene. Construction of plasmids:

Constructs containing the GFP reporter gene were made by excising the lacZ gene from CP2615 (nFus1-lacz AmpR TRP1 2mu) with SalI and EagI and replacing it with the wild type GFP gene (Clontech) to produce Cp2759 (nFus1-GFPwt 2mu AmpR TRP1).

Cp2759 (nFus1-GFPwt 2mu AmpR TRP1) was used as a template to introduce mutations into the wildtype gene (Cp2759) using the quick-change mutagenesis protocol (Stratagene). The resulting plasmid, Cp4567 (nFus1-GFP F100S M154T V164A 2mu AmpR TRP1) contains a threonine at position 154 which was changed to a yeast codon biased threonine. The mutations are the same as those described by Cramer *et al.*, Nature Biotech 1996 14:315-319.

Cp5681 (nfus1p-GFPS65T 2mu AmpR TRP1) was made by digesting the Cp5600 with sites internal to GFP gene using NcoI and BstBI. The 450 base pair fragment was ligated into Cp4567 digested with the same enzymes. The Cp5681 was digested with BglII and EagI and the 1.1 kb fragment containing nFus1p-GFPS65T was subcloned into the multiple cloning site of following yeast integrative vectors; Cp1007 (HIS3 AmpR); Cp1009 (TRP1 AmpR); Cp1010 (LEU2 AmpR) and; Cp1011 (URA3 AmpR) that were digested with BamHI/EagI to generate: Cp5772 (nFus1GFP65T

- 109 -

HIS3 AmpR); Cp5773 (nFus1GFPs65T TRP1 AmpR); Cp5774 (nFus1GFPs65T LEU2 AmpR); Cp5775 ((nFus1GFPs65T LEU2 AmpR).

Yeast strains:

5 Cp5774 and Cp5775 were linearized with BbsI, which cuts within the LEU2 gene, and were transformed into CY1829 (bar1::hisG fus1-HIS3 tbt1-1 ura3 trp1 leu2 his3 SUC+) and plated on URA LEU minus plates, yielding three transformants. The three transformants were tested for GFP fluorescence via microscopy and fluorometer and the candidate with the highest fluorescence was designated, CY 16363  
10 (bar1::hisG fus1-HIS3 tbt1-1 ura3::URA3::fus1-GFPS65T trp1 leu2::LEU::fus1-GFPS65T his3 SUC+).

Similarly, Cp5773 was linearized with Bsu36I, and Cp5775 was linearized with BbsI, and both transformed simultaneously into CY10981 (GPA+sst2\*2far1\*1442tbt1-1fus1-HIS3 can1 ste14::trp1::LYS2 ste3\*1156 lys2 ura3  
15 leu2 trp1 his3). 10 transformants were isolated and the functional expression of GFPs65T was tested by transforming mammalian receptor, ML1aR (Cp2695). The fluorescence was observed by microscope and by FACs or a fluorometer. The candidate with the highest fluorescence was designated, CY16637.

20 For TRP1 directed integration of fus1-GFPS65T, Cp5773 was linearized with Bsu36I, which cuts within TRP1 gene, and transformed into CY10981. Transformants were isolated and functional expression of GFPs65T was tested by transforming mammalian receptor. The fluorescence was observed by microscope and by FACs or a fluorometer. The candidate with the highest fluorescence was designated,  
25 CY16638.

Standard Green Fluorescent Protein (GFP) assay

Yeast strains were grown overnight at 30°C to stationary phase in the appropriate media, as described above. The strains were diluted to 0.2 OD/ml and  
30 allowed to grow until an optical density of OD = 0.4-0.5 was attained. Ligand was added to the cells which were induced for 2 ½ hrs. Yeast cells were harvested at approximately 1.0-1.2 OD/ml at which time the cells show the highest fluorescence.

For measurements are taken using fluorometer using GFP:

35 The dose response for Ste2 was investigated using CY16363 which was grown overnight in YEPD at 30°C. After overnight culture, the yeast cells were diluted

to 0.25 OD/ml and allowed to grow for an additional 2 ½ hrs. GFP fluorescence was determined with increasing concentrations of  $\alpha$ Factor.  $\alpha$ Factor was added to 6 mls YEPD (in triplicate) at the following concentrations: 1pM, 10pM, 100pM, 1nM, 10nM, 100nM, 1 $\mu$ M. Cultures were incubated at 30°C with ligand for 2hrs. Cells were  
5 harvested at 1.1-1.5 OD/ml. Yeast cells at 6.5 ODs ( $6.5 \times 10^7$  cells/ml) were centrifuged and washed once with sterile water. The yeast cells were resuspended in 100 $\mu$ l of sterile water and transferred to 96 well plate to be read by fluorometer at excitation wavelength 485nm, an emission wavelength of 535 nm and constant gain of 50.

10 For measurements are taken using a FACS machine and GFP.

The dose response for Ste2 was investigated using CY16363 which was grown overnight in YEPD 30°C. After overnight culture, the cells were diluted to an OD<sub>600</sub> of 0.2 OD/ml. The yeast cells were grown for and additional 2 ½ hrs or until an OD<sub>600</sub> 0.4-0.6 OD/ml was attained.  $\alpha$ Factor was added to 6 mls YEPD (in triplicate) at  
15 the following concentrations: 1pM, 10pM, 100pM, 1nM, 10nM, 100nM, 1 $\mu$ M. Cultures were incubated at 30°C with ligand for 2hrs. The yeast cells were induced with ligand for 2 ½ hrs at 30°C until OD<sub>600</sub> is about 0.8-1.0 OD/ml. Yeast cells were harvested and washed once with water. The yeast cells were resuspended in water to a final concentration OD.0.1D/ml. Typically, 10,000-50,000 cells were used for the  
20 analysis using standard FACS analysis procedures.

For measurements using the LacZ assay

For comparative purposes, the dose response for Ste2 was investigated using the lacZ assay. CY16363 was grown overnight in YEPD at 30°C. After overnight culture, the cells were diluted to an OD<sub>600</sub> of 0.2 OD/ml. The yeast cells were grown for and additional 2 ½ hrs or until an OD<sub>600</sub> 0.5 OD/ml was attained. The yeast cells were diluted to an 0.1 OD/ml and 100 $\mu$ l of the cells were transferred to a 96 well plate. 1 $\mu$ l of  $\alpha$ Factor was added to each well to produce final concentrations as follows: 0pM, 1pM, 10pM, 100pM, 1nM, 10nM, 100nM, 1 $\mu$ M. Yeast cells were induced with ligand for 3 hrs at 30°C. 20 $\mu$ l Triton/CPRG substrate was added to each well and incubated at 30°C for 25 min. and 45 min.

Analysis of Ste2 using fus1-GFP65T by both FACS and fluorometric analysis demonstrates a significant improvement in the sensitivity compared with and fus1-Lac Z.

35

Dose Response for ML1a receptor measured using a fluorometer and GFP

The dose response for the ML1a receptor was investigated using CY16639 which was grown overnight in LEU minus media at 30°C. After overnight culture, the cells were diluted to an OD<sub>600</sub> of 0.25 OD/ml into 6 ml LEU6.8. Cells were grown an additional 2 ½ hrs (OD = 0.4-0.5). ML1a ligand was added in increasing concentrations as follows: 0pM, 10pM, 100pM, 1nM, 10nM, 100nM, 1μM. Cells were induced with the ligand for 2 ½ hrs at 30°C. Yeast cells at a density of 6 x 10<sup>7</sup> cells were harvested and washed once with water. The yeast cells were resuspended in water and transferred to 96-well plate to be read by fluorometer at excitation wavelength 485nm, an emission wavelength of 535 nm and constant gain of 50.

10

Dose Response for ML1a receptor measured using a FACS machine and GFP

The dose response for the ML1a receptor was investigated using CY16639 which was grown overnight in LEU minus media at 30°C. After overnight culture, the cells were diluted to 0.2 OD/ml into 2ml LEU6.8. ML1a ligand was added at the same concentrations as the fluorometer assay, described above. Cells were induced with the ligand for 2 ½ hrs at 30°C (OD = 0.7-1.2). The cells were harvested and washed once with water. The cells were resuspended in sterile water to a final OD<sub>600</sub> = 0.1 OD/ml and analysed using standard FACS analysis.

20 Dose Response for ML1a receptor measured using the LacZ assay

For comparative purposes, the dose response for ML1a receptor was investigated using the lacZ assay. CY15438 (GPA1+3907 sst2\*far1\*1442 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3\*1156 lys2 ura3 leu2 trp1 his3) was grown overnight at 30°C. After overnight culture, the cells were diluted to 0.2 OD/ml. Cells were grown for an additional 3 hrs to 0.4 OD/ml. The cells were then diluted to 0.1 OD/ml and 100 μl transferred to 96-well plate. 1.2μl of the ligand was added, such that the concentration of DMSO was 0.24% for each well. The ML1a receptor was added at the following concentrations: 0pM, 1 pM, 10pM, 100pM, 1nM, 10nM, 100nM, 10μM, 100μM. The cells were incubated at 30°C for 3 hrs with ligand. 20μl of Triton/CPRG was added to each well and incubated at 37°C for 1 hr. 60μl 1M Na<sub>2</sub>CO<sub>3</sub> was added to stop the reaction and the absorbance was measured at OD<sub>570</sub>.

Two variants of the GFP wild-type were tested and compared to nFus1-GFPwt (Cp2759). the variants GFP F100S M154T V164A (Cp4567) and GFP S65T (Cp5681) both show higher relative fluorescence compared to the wildtype GFP. GFP S65T showed the highest fluorescence. GFP reporter genes can be used to improve sensitivity which can be used to control for visualizing uniformity or variation in the

signal within a population of cells. Analysis of ML1a GPA fus1-GFP65T by both FACs and fluorometric analysis demonstrates a significant improvement in the sensitivity compared with and fus1-Lac Z.

- 113 -

SEQUENCE LISTING

5 (1) GENERAL INFORMATION:

(i) APPLICANT:

10 (ii) TITLE OF INVENTION: DNA Sequences

(iii) NUMBER OF SEQUENCES:

15 (iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: LAHIVE & COCKFIELD  
(B) STREET: 28 State Street  
(C) CITY: Boston  
(D) STATE: Massachusetts  
(E) COUNTRY: USA  
20 (F) ZIP: 02109-1875

25 (v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

30 (vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:  
(B) FILING DATE:  
30 (C) CLASSIFICATION:

35 (vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER:  
(B) FILING DATE:

35 (viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Catherine J. Kara  
(B) REGISTRATION NUMBER: 41,106  
40 (C) REFERENCE/DOCKET NUMBER: CPI-068

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (617) 227-7400  
(B) TELEFAX: (617) 742-4214

45

(2) INFORMATION FOR SEQ ID NO:1:

50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

- 114 -

(v) FRAGMENT TYPE: internal

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Leu Leu Leu Leu Gly Ala Gly Glu Ser Gly  
1 5 10

10

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

20

(v) FRAGMENT TYPE: internal

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gly Ser Gly Glu Ser Gly Asp Ser Thr  
1 5

30

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

45

Gln Ala Arg Lys Leu Gly Ile Gln  
1 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

50

- 115 -

(v) FRAGMENT TYPE: internal

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu Ile His Glu Asp Ile Ala Lys Ala  
1 5

10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

20

(v) FRAGMENT TYPE: internal

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:  
Asp Val Gly Gly Gln  
1 5

30

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: cDNA

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGATCTAGAG GATCCTGGTA CGTTCCTC

28

45

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: cDNA

- 116 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

5           GTCGCTAGCC AAGCTTGCAT GCCTGCAG  
28

10           (2) INFORMATION FOR SEQ ID NO:8:

10           (i) SEQUENCE CHARACTERISTICS:  
15           (A) LENGTH: 32 base pairs  
              (B) TYPE: nucleic acid  
              (C) STRANDEDNESS: single  
              (D) TOPOLOGY: linear

20           (ii) MOLECULE TYPE: cDNA

20           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

25           GCAGTCATGA GATTTCCCTTC AATTTTTACT GC  
32

25           (2) INFORMATION FOR SEQ ID NO:9:

30           (i) SEQUENCE CHARACTERISTICS:  
35           (A) LENGTH: 30 base pairs  
              (B) TYPE: nucleic acid  
              (C) STRANDEDNESS: single  
              (D) TOPOLOGY: linear

35           (ii) MOLECULE TYPE: cDNA

40           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

40           CAGCCCATGG CTTCAGCCTC TCTTTTATCC  
30

45           (2) INFORMATION FOR SEQ ID NO:10:

45           (i) SEQUENCE CHARACTERISTICS:  
50           (A) LENGTH: 41 base pairs  
              (B) TYPE: nucleic acid  
              (C) STRANDEDNESS: single  
              (D) TOPOLOGY: linear

50           (ii) MOLECULE TYPE: cDNA

- 117 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GATCCGCTTA CCCATACGAT GTTCCAGATT ACGCTGCTTG A  
41

5

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
10 (A) LENGTH: 31 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

20 GATCTCAAGC AGCGTAATCT GGAACATCGT A  
31

(2) INFORMATION FOR SEQ ID NO:12:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 12 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ser Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ala  
1 5 10

40

(2) INFORMATION FOR SEQ ID NO:13:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 38 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

- 118 -

GATATATTAA GGTAGGAAAC CATGGGGTGT ACAGTGAG  
38

(2) INFORMATION FOR SEQ ID NO:14:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: cDNA

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGAGGGCTCGA GGGAACGTAT AATTAAAGTA GTG

33

20

(2) INFORMATION FOR SEQ ID NO:15:

25

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30

(ii) MOLECULE TYPE: cDNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

34

(2) INFORMATION FOR SEQ ID NO:16:

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: cDNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCCGAATCCA CCAATTCTT TACG

24

- 119 -

(2) INFORMATION FOR SEQ ID NO:17:

- 5           (i) SEQUENCE CHARACTERISTICS:  
              (A) LENGTH: 27 base pairs  
              (B) TYPE: nucleic acid  
              (C) STRANDEDNESS: single  
              (D) TOPOLOGY: linear

10           (ii) MOLECULE TYPE: cDNA

              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

15           GCGGCGTCGA CGCGGCCGCG TAACAGT  
27

(2) INFORMATION FOR SEQ ID NO:18:

- 20           (i) SEQUENCE CHARACTERISTICS:  
              (A) LENGTH: 37 base pairs  
              (B) TYPE: nucleic acid  
              (C) STRANDEDNESS: single  
              (D) TOPOLOGY: linear

25           (ii) MOLECULE TYPE: cDNA

30           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

              CTGCTGGAGC TCCGCCTGCT GCTGCTGGGT GCTGGAG  
37

35

(2) INFORMATION FOR SEQ ID NO:19:

- 40           (i) SEQUENCE CHARACTERISTICS:  
              (A) LENGTH: 43 base pairs  
              (B) TYPE: nucleic acid  
              (C) STRANDEDNESS: single  
              (D) TOPOLOGY: linear

45           (ii) MOLECULE TYPE: cDNA

              (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

50           CTGCTGGTCG ACGCGGCCGC GGGGGTTCCCT TCTTAGAAGC AGC  
43

(2) INFORMATION FOR SEQ ID NO:20:

- 120 -

- 5                   (i) SEQUENCE CHARACTERISTICS:  
                     (A) LENGTH: 30 base pairs  
                     (B) TYPE: nucleic acid  
                     (C) STRANDEDNESS: single  
                     (D) TOPOLOGY: linear

10                  (ii) MOLECULE TYPE: cDNA

15                  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

                 GGGCTCGAGC CTTCTTAGAG CAGCTCGTAC  
15                  30

20                  (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
                     (A) LENGTH: 37 base pairs  
                     (B) TYPE: nucleic acid  
                     (C) STRANDEDNESS: single  
                     (D) TOPOLOGY: linear

25                  (ii) MOLECULE TYPE: cDNA

30                  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

                 CTGCTGGAGC TCAAGTTGCT GCTGTTGGGT GCTGGGG  
37

35                  (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
                     (A) LENGTH: 44 base pairs  
                     (B) TYPE: nucleic acid  
                     (C) STRANDEDNESS: single  
                     (D) TOPOLOGY: linear

40                  (ii) MOLECULE TYPE: cDNA

45                  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

                 CTGCTGGTCG ACGCGGCCGC GCCCCTCAGA AGAGGCCGCG GTCC  
44

50                  (2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:  
                     (A) LENGTH: 29 base pairs

- 121 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: cDNA

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGGCTCGAGC CTCAGAAGAG GCCGCAGTC  
29

15 (2) INFORMATION FOR SEQ ID NO:24:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: cDNA

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CTGCTGGAGC TCAAGCTGCT GCTACTCGGT GCTGGAG  
37

30 (2) INFORMATION FOR SEQ ID NO:25:

35 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

45 CTGCTGGTCG ACGCGGCCGC CACTAACATC CATGCTTCTC AATAAGTC  
49

(2) INFORMATION FOR SEQ ID NO:26:

50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- 122 -

(ii) MOLECULE TYPE: cDNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGGCTCGAGC ATGCTTCTCA ATAAAGTCCA C

31

10

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

25

GCATCCATCA ATAATCCAG

19

(2) INFORMATION FOR SEQ ID NO:28:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: cDNA

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GAAACAATGG ATCCACTTCT TAC

23

45

GYPVAGSNHFIEGLKNAQKNSQMSNSNSVCCTL M (SEQ ID NO: 29)

50 DPLLTPVPASENPFREKKFFCAIL (SEQ ID NO:30)

GYPVAGSNHFIEGLKNAQKNSQMSNSNSVCAIL (SEQ ID NO:31)

GYPVAGSNHFIEGLKFFCCTL M (SEQ ID NO:32)

- 123 -

5 GYPVAGSNHFIEGLKNPFREKKFFCCTL  
GYPVASENPFREKKFFCCTL  
GYPVAGSNPFREKKFFCCTL  
GYPVAGSNHFREKKFFCCTL  
10 GYPVAGSNHFIEKKFFCCTL  
GYPVAGSNHFIEGLKNPKELKGGCCTL  
GYPVAGSNHFIEGLKNPFRPQKVCCCTL  
15 GYPVAGSNHFIEGLKNPKDKKPCCCTL  
GYPVAGSNHFIEGLKNPFREPRSCCTL  
GYPVAGSNHFIEGLKNPFKEKGSCCTL  
20 GYPVAGSNHFIEGLKNAQKNPFREKKFFCCTL  
GYPVAGSNHFIEGLKNPFREKKFFCCTL  
GYPVAGSNHFIEGLKNPFREKKFFCCTL  
25 CCCCTCGAGTTCCCTTCTTAGAGCAGCT (SEQ ID NO:45)  
CCCGAAGACCAAGCTTGACCAGGTTATTATA (SEQ ID NO:46)  
30 AAGGAAGACTTAGCTTATAATTTGGGCTTAGTT (SEQ ID NO:47)  
TTT CTT GTC ACT CCG TTT CTA AC (SEQ ID NO 48)  
CCCCGTCTCAAGAGCGCTGGCCGCCACATC (SEQ ID NO:49)  
35 CCCCCGTCTCACTCTGAACGCCGCAAGTGGATCC (SEQ ID NO:50)  
AGCAAGCAGATCTTGCTTGTG (SEQ ID NO:51)  
40 GATGTGGCGGCCAGCGCTCTGAACGCCGCAAGTGGATC (SEQ ID NO:52)  
DVGGQRSERRKWI (SEQ ID NO:53)  
TCGTCTGGAGCTCAAGCTGCTGCTTTGGGCCAGGCAGAGCGGGAAAGAGC (SEQ ID NO:54)  
45 CTGCTGGTCGACGCCGCCGGTCACAGCAGGTTGAT CTCGTCCAG (SEQ ID NO:55)  
GCTAGGTCTCAACCAGTCATCCTCTTCTAACAAAACC (SEQ ID NO:56)  
50 ACGTGGTCTCATGGTGTGCTTTGAACCAGGGTAGT (SEQ ID NO:57)  
AGCGGCTGCAGATTCCATTC (SEQ ID NO:58)  
CTACCCCTGGTTCAAAAGCACACCAGTCATCCTCTTCTAACAAA (SEQ ID NO:59)

- 124 -

LPWFKSTPVILFLNK (SEQ ID NO:60)  
GGCGTCTCACATGGGATGTACGCTGAGCG (SEQ ID NO:61)  
5 GGGGTCGACTCAGTAGAGTCAACAGCCC (SEQ ID NO:62)  
CCCCCGTCTCATAGGGTATCAATGAAAAACTTGTGTTA (SEQ ID NO:63)  
10 CCCCCGTCTCACCTATCATCCTCTTCCTCAACAAG (SEQ ID NO:64)  
AACAAACAAGTTTCATTGATAACCCCTATCATCCTCTTCCTCAACAAG (SEQ ID NO:65)  
NNKFIDTPIILFLNK (SEQ ID NO:66)  
15 AAAAGAGCCAATGATGTCATCGAGCAATCGTTGCAGCTGGAGAAACAACGTGACAAGAATGAGCT  
(SEQ ID NO:67)  
CATTCTGTCACGTTCTCCAGCTGCAACGATTGCTCGATGACATCATTGGCTCTTTGTT  
(SEQ ID NO:68)  
20 CATGGGGTGTACAGTGAGTACGCAAACAATAGGAGATGAAAGTGATCCTTCTGCAGAAC (SEQ ID  
NO:69)  
YGCAGAAAAGGATCACTTCATCTCCTATTGTTGCGTACTCACTGTACACCC (SEQ ID NO:70)  
25 CTTCTCAACGTCCCCATCATCCTC (SEQ ID NO:71)  
GAGGATGATGGGGACGTTGAAGAAC (SEQ ID NO:72)  
GGATGTGGCGCCAGAGGTACAG (SEQ ID NO:73)  
30 CTGTGACCTCTGGCGCCACATCC (SEQ ID NO:74)  
GCGGAGCTCMNNMNNMNNMNNCTTTCTAATTGCAAGGATTGTCGATAACGTCTAGCTCTTAT  
TCTGCAGGG (SEQ ID NO:75)  
35 CCCTGCAGATAAGAGAGCTAATGACGTTATCGAACAACTCTGCAATTA (SEQ ID NO:76)  
  
GCATCACATCAATAATCCAG (SEQ ID NO:77)  
40 AACCCGGAACGATTAACGAGATCAAGAAC (SEQ ID NO:78)  
LEKQRDKNE (SEQ ID NO:79)  
45 GxGxxG (SEQ ID NO:80)  
MGXaaS (SEQ ID NO:82)  
50 CAAAX (SEQ ID NO:83)  
ACC CGG AAC GAT TTA ACG AG (SEQ ID NO:84)  
GAT TGG AGC CGG TGA CTA CC (SEQ ID NO: 85)

- 125 -

KREAEA (SEQ ID NO: 86)

TTT GAA GGG CCG TAT AAA GAC (SEQ ID NO: 87)

5

ACG TCT CGA GGT CGA CTC ATA GAA GCT CAT ATT GTT TAA GG  
(SEQ ID NO: 88)

10

ACG TCT CGA GAT CGA CTC AAA ATA GAC CAC AGT CTT TAA GGT TTT  
GCT GGA TG (SEQ ID NO: 89)

ACG TCT CGA GGT CGA CTC AAC AAA GAC CAA TAT ATT TAA GGT TTT  
GCT GGA TG (SEQ ID NO: 90)

15

ACG TCT CGA GGT CGA CTC ATA CAA GAT TAT ATT CTT TAA GGT TTT  
GCT GGA TG (SEQ ID NO: 91)

ACG TCT CGA GGT CGA CTC ATT GAA GCA TAA TGT CTT TAA GGT TTT  
GCT GGA TG (SEQ ID NO: 92)

20

ACG TCT CGA GGT CGA CTC ATT GAA GCA TAA GTT GTT TAA GGT TTT  
GCT GGA TG (SEQ ID NO: 93)

25

ACG TCT CGA GGT CGA CTC AAA GAA GAT TAA TTT CAT CAA GGT TTT  
GCT GGA TGA TT (SEQ ID NO: 94)

ACG TCT CGA GGT CGA CTC ATA GAA GCT CAT ATT GTC TAA GGT TTT  
GCT GGA TG (SEQ ID NO: 95)

30

ACG TCT CGA GAT CGA CTC AAT ATA GAC CAC ATC CTT TAA GGT TTT  
GCT GG (SEQ ID NO: 96)

AVTDLIIQQNLKKIGII (SEQ ID NO: 97)

AVTDVIIKNNLKDCGLF (SEQ ID NO: 98)

35

AVKDITQLNLKEYNLV (SEQ ID NO: 99)

AVKDITQENLKDIMLQ (SEQ ID NO: 100)

DCRDIIQRMHLRQYELL (SEQ ID NO: 101)

DVKDITLHDNLKQLMLQ (SEQ ID NO: 102)

AVTDVIIQNNLKYGGLC (SEQ ID NO: 103)

40

DVRDSVLARYLDEINLL (SEQ ID NO: 104)

DCRDIIQRMHLRQYELL (SEQ ID NO: 105)

AVTDIIIANNLRGCGLY (SEQ ID NO: 106)

**GPA1-Gαi2(5):**

45

MGCTVSTQTIGDESDPFLQNKRANDVIEQLSQLKEQRDKNEIKLLLGGAGESGKSTVLKQLKLLHQGGFSHQERLQYAQVIWADAIQSMKILIIQARKLGQLQDCDDPINNKDLFACKRILLKAKALDYINASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAAFDEDGNISNVKSDTDRAETVTQNEDADRNNSSRINLQDICLKDLNQEGDDQMFVRKTSREIQGQNRNLIHEDIAKAIKQLWNNNDKGKIKQCFARSNEFQLEGSAAYFDNIEKFASPNYVCTDEDILKGRIKTTGITETEFNIGSSKFKVLDAGGQRSERKKWIHCPEGITAVLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLFEEKVKSMPIRKYFPDYQGRVGDAEAGLKYEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVTDLIIQQNLKDCGLF (SEQ ID NO: 107)

**GPA1-Gαq(5):**

- 126 -

- MGCTVSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLLLLH  
 QGGFQSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYIN  
 ASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRAETVTQNEDAD  
 RNNSSRINLQDICKDLNQEGDDQMFVRKTSREIQGQNRRNLIHEDIAKAIQLWNNDKGIKQCFA  
 5 RSNEFQLEGSAAYFDNIEKFASPVYVCTDEDILKGRIKTTGITETEFNIGSSKFVLDAGGQRSER  
 KKWIHCSEGITA VLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLF  
 EEKVKSMPIRKYFPDYQGRVGDAEAGLKYEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVT  
 DLIIQQNLKEYNLV (SEQ ID NO: 108)
- 10 **GPA1-G<sub>12</sub>(5):**  
 MGCTVSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLLLLH  
 QGGFQSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYIN  
 ASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRAETVTQNEDAD  
 RNNSSRINLQDICKDLNQEGDDQMFVRKTSREIQGQNRRNLIHEDIAKAIQLWNNDKGIKQCFA  
 15 RSNEFQLEGSAAYFDNIEKFASPVYVCTDEDILKGRIKTTGITETEFNIGSSKFVLDAGGQRSER  
 KKWIHCSEGITA VLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLF  
 EEKVKSMPIRKYFPDYQGRVGDAEAGLKYEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVT  
 DLIIQQNLKDMLQ (SEQ ID NO: 109)
- 20 **GPA1-G<sub>13</sub>(5):**  
 MGCTVSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLLLLH  
 QGGFQSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYIN  
 ASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRAETVTQNEDAD  
 RNNSSRINLQDICKDLNQEGDDQMFVRKTSREIQGQNRRNLIHEDIAKAIQLWNNDKGIKQCFA  
 25 RSNEFQLEGSAAYFDNIEKFASPVYVCTDEDILKGRIKTTGITETEFNIGSSKFVLDAGGQRSER  
 KKWIHCSEGITA VLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLF  
 EEKVKSMPIRKYFPDYQGRVGDAEAGLKYEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVT  
 DLIIQQNLKQLMLQ (SEQ ID NO: 111)
- 30 **GPA1-G<sub>az</sub>(5):**  
 MGCTVSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLLLLH  
 QGGFQSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYIN  
 ASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRAETVTQNEDAD  
 RNNSSRINLQDICKDLNQEGDDQMFVRKTSREIQGQNRRNLIHEDIAKAIQLWNNDKGIKQCFA  
 35 RSNEFQLEGSAAYFDNIEKFASPVYVCTDEDILKGRIKTTGITETEFNIGSSKFVLDAGGQRSER  
 KKWIHCSEGITA VLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLF  
 EEKVKSMPIRKYFPDYQGRVGDAEAGLKYEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVT  
 DLIIQQNLKYIGLC (SEQ ID NO: 112)
- 40 **GPA1-G<sub>az</sub>(5):**  
 MGCTVSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLLLLH  
 QGGFQSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYIN  
 ASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRAETVTQNEDAD  
 RNNSSRINLQDICKDLNQEGDDQMFVRKTSREIQGQNRRNLIHEDIAKAIQLWNNDKGIKQCFA  
 45 RSNEFQLEGSAAYFDNIEKFASPVYVCTDEDILKGRIKTTGITETEFNIGSSKFVLDAGGQRSER  
 KKWIHCSEGITA VLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLF  
 EEKVKSMPIRKYFPDYQGRVGDAEAGLKYEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVT  
 DLIIQQNLQYELL(SEQ ID NO: 110)
- 50 **GPA1-G<sub>az</sub>(6):**  
 MGCTVSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLLLLH  
 QGGFQSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYIN  
 ASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRAETVTQNEDAD  
 RNNSSRINLQDICKDLNQEGDDQMFVRKTSREIQGQNRRNLIHEDIAKAIQLWNNDKGIKQCFA

- 127 -

RSNEFQLEGSAAYYFDNIEKFASPNYVCTDEDILKGRIKTTGITETEFNIGSSKFKVLDAGGQRSER  
 KKWIHCSEGITAFLVFLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFLNLKIDLF  
 EEKVKSMPIRKYFPDYQGRVGDAEAGLKYFEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVT  
 DLIIQQNLRQYELL(SEQ ID NO: 115)

5

**GPA1-G<sub>α16</sub>(6):**

MGCTVSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLKLLH  
 QGGFSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYIN  
 ASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRDAETVTQNEDAD  
 10 RNNSSRINLQDICKDLNQEGDDQMFRKTSREIQGQNRRNLIHEDIAKAIKQLWNNNDKGIKQCFA  
 RSNEFQLEGSAAYYFDNIEKFASPNYVCTDEDILKGRIKTTGITETEFNIGSSKFKVLDAGGQRSER  
 KKWIHCSEGITAFLVFLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFLNLKIDLF  
 EEKVKSMPIRKYFPDYQGRVGDAEAGLKYFEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVT  
 DLIIQQNLDEINLL(SEQ ID NO: 114)

15

**GPA1-G<sub>α0</sub>(5):**

MGCTVSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLKLLH  
 QGGFSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYIN  
 ASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRDAETVTQNEDAD  
 20 RNNSSRINLQDICKDLNQEGDDQMFRKTSREIQGQNRRNLIHEDIAKAIKQLWNNNDKGIKQCFA  
 RSNEFQLEGSAAYYFDNIEKFASPNYVCTDEDILKGRIKTTGITETEFNIGSSKFKVLDAGGQRSER  
 KKWIHCSEGITAFLVFLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFLNLKIDLF  
 EEKVKSMPIRKYFPDYQGRVGDAEAGLKYFEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVT  
 DLIIQQNLKGCGLY(SEQ ID NO: 113)

25

ACGTGGTCTCCATGACTTGGAACTATTATGGCTTGTCTTAGTACGAAACAATAGGA  
 GACG (SEQ ID NO: 116)

GTATTTGAACCACTTAGAG (SEQ ID NO: 117)

30

**G<sub>αq</sub>(1-11)-GPA1(6-467)-G<sub>αq</sub>(355-359):**

MTLESIMACCLSTQTIGDESDPFLQNKRANDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQ  
 LKLLHQGGFSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKA  
 LDYINASVAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRDAETVTQ  
 35 NEDADRNNSSRINLQDICKDLNQEGDDQMFRKTSREIQGQNRRNLIHEDIAKAIKQLWNNNDKGI  
 KQCFARSNEFQLEGSAAYYFDNIEKFASPNYVCTDEDILKGRIKTTGITETEFNIGSSKFKVLDAGG  
 QRSERKKWIHCSEGITAFLVFLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFLFL  
 NKIDLFEEKVKSMPIRKYFPDYQGRVGDAEAGLKYFEKIFLSLNKTNKPIYVKRTCATDTQTMKF  
 VLSAVTDLIIQQNLKEYNLV(SEQ ID NO: 118).

40

GTCTAAAATGAAGAGGAGATAGTAG (SEQ ID NO: 119)

GATCCGTCTCACTTCAGAAAGACAACAAGCCATAATAG (SEQ ID NO: 120)

45

GATCCGTCTCTGAAGAACGCTAAGGAGGCTAGAAGAATTAATGATGTCATCGAGCAATCGTTG  
 C (SEQ ID NO: 121)

GTATTTGAACCACTTAGAG (SEQ ID NO: 122)

50

**G<sub>αq</sub>(1-21)-GPA1(24-467)-G<sub>αq</sub>(355-359):**

MTLESIMACCLSEEAKEARRINDVIEQLQLEKQRDKNEIKLLLLGAGESGKSTVLKQLKLLHQG  
 GFSHQERLQYAQVIWADAIQSMKILIIQARKLGQLDCDDPINNKDLFACKRILLKAKALDYINAS  
 VAGGSDFLNDYVLKYSERYETRRRVQSTGRAKAASFDEDGNISNVKSDTDRDAETVTQNEDADR

**- 128 -**

NNSSRINLQDICKDLNQEQQMFVRKTSREIQGQNRRNLIHEDIAKAIQLWNNDKGIKQCFAR  
SNEFQLEGSAAYYFDNIEKFASPNYVCTDEDILGRIKTTGITETEFNIGSSKFKVLDAGGQRSERK  
KWIHCSEGITAVLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLFE  
EKVKSMPIRKYFPDYQGRVGDAEAGLKYEKIFLSLNKTNKPIYVKRTCATDTQTMKFVLSAVTD  
5 LIIQQNLKEYNLV(SEQ ID NO: 123).

**Claims**

1. A recombinant yeast cell which comprises:
  - a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast pheromone receptor in a pheromone response pathway of the yeast cell; and
  - a chimeric G protein subunit selected from the group consisting of:
    - 10 a non-naturally occurring G protein subunit which comprises a sequence from a heterologous G protein subunit in which at least one amino acid substitution has been introduced compared to the wild type sequence;
    - an endogenous STE 18 subunit operably linked to a polypeptide of a heterologous Gy subunit;
    - 15 an endogenous yeast Gpa1 subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit; and
    - an endogenous yeast Gpa1 subunit in which at least the last four C-terminal amino acids of said Gpa1 are replaced with at least the last four C-terminal amino acids of a first heterologous G protein subunit, and in which the N-terminus of said Gpa1 is operably linked to at least the first five N-terminal amino acids of a second heterologous G protein subunit, wherein said first and second heterologous G protein subunits are the same or different;
    - such that expression of said chimeric G protein subunit functionally integrates said 25 heterologous GPCR into the pheromone response pathway of said yeast cell; and wherein modulation of the signal transduction activity of said heterologous GPCR by an extracellular signal provides a detectable signal.

2. The yeast cell of claim 1, wherein said chimeric G protein subunit comprises an endogenous yeast G<sub>α</sub>1 subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit.

3. The yeast cell of claim 1, wherein said chimeric G protein subunit comprises an endogenous yeast G<sub>α</sub>1 subunit in which at least the last four C-terminal amino acids of said G<sub>α</sub>1 are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, and in which the N-terminus of said G<sub>α</sub>1 is operably linked to at least the first five N-terminal amino acids of a heterologous G protein subunit, wherein said heterologous G protein subunits are the same or different.

4. The yeast cell of claim 1, wherein said chimeric G protein subunit comprises a non-naturally occurring G protein subunit which comprises a sequence from a heterologous G protein subunit in which at least one amino acid substitution has been introduced compared to the wild type sequence.

5. The yeast cell of claim 1, wherein said chimeric G protein subunit comprises an endogenous STE 18 subunit operably linked to a polypeptide of a heterologous G<sub>γ</sub> subunit.

6. The yeast cell of claim 4, wherein said non-naturally occurring G protein subunit is a mutant mammalian G<sub>α</sub> subunit selected from the group consisting of: G<sub>α</sub> 16(S270P); G<sub>α</sub>s(D229S); G<sub>α</sub>s(D229V); G<sub>α</sub>s(N254D); G<sub>α</sub>s(S286P); G<sub>α</sub>s (E10K); G<sub>α</sub>i2-G<sub>α</sub>oB (S280P); G<sub>α</sub>12 (Q229L); G<sub>α</sub>12 (G228A); and G<sub>α</sub>i2 (S288P).

7. The yeast cell of claim 4, wherein said non-naturally occurring G protein subunit is a yeast-mammalian G protein subunit chimera comprising a first polypeptide from a yeast G protein subunit and a second polypeptide from a mutant mammalian G protein subunit.

5

8. The yeast cell of claim 7, wherein said second polypeptide comprises a mutant mammalian G $\alpha$  subunit selected from the group consisting of: G $\alpha$ 16(S270P); G $\alpha$ s(D229S); G $\alpha$ s(D229V); G $\alpha$ s(N254D); G $\alpha$ s(S286P); G $\alpha$ s (E10K); G $\alpha$ i2-G $\alpha$ oB (S280P); G $\alpha$ 12 (Q229L); G $\alpha$ 12 (G228A); and G $\alpha$ i2 (S288P).

10

9. A yeast cell comprising a chimeric G protein subunit, said chimeric G protein subunit comprising a first polypeptide from a yeast G protein subunit and a second polypeptide from a heterologous G protein subunit, wherein said first polypeptide is a polypeptide from yeast STE 18.

15

10. The yeast cell of claim 9, wherein said first polypeptide is from yeast STE 18 and said second polypeptide is from a heterologous G protein  $\gamma$  subunit.

20

11. The yeast cell of claim 10, wherein said heterologous G protein subunit is mammalian.

12. The yeast cell of claim 11, wherein said heterologous G protein subunit is human.

25

13. The yeast cell of claim 12, wherein at least one of said first and second polypeptides comprises a naturally occurring amino acid sequence.

14. The yeast cell of claim 12, wherein at least one of said first and second polypeptides comprises a non-naturally occurring amino acid sequence.

15. The yeast cell of claim 9, further comprising a heterologous G protein  
5 coupled receptor, which receptor is functionally integrated into the yeast cell.

16. The yeast cell of claim 15, wherein said chimeric G protein subunit demonstrates enhanced coupling to the heterologous G protein coupled receptor when compared to that demonstrated by an endogenous yeast G protein subunit.

10

17. The yeast cell of claim 10, wherein said second polypeptide is derived from the human Gy2 subunit.

15

18. The yeast cell of claim 17, wherein said second polypeptide comprises the amino acid sequence Arg Glu Lys Lys Phe Phe (amino acids 19-24 of SEQ ID NO: 33).

19. The yeast cell of claim 18, wherein said second polypeptide comprises the sequence shown in SEQ ID NO: 33.

20

20. The yeast cell of claim 9, further comprising a second chimeric G protein subunit, said second chimeric G protein subunit comprising a first polypeptide from a yeast G protein subunit and a second polypeptide from a mammalian G protein subunit, wherein said second chimeric G protein subunit is different from said first chimeric G protein subunit.

21. The yeast cell of claim 20, wherein said second polypeptide of said second chimeric G protein subunit is from a protein selected from the group consisting of: a mammalian G $\alpha$  subunit, a mammalian G $\beta$  subunit, and a mammalian G $\gamma$  subunit.

5 22. The yeast cell of claim 15, wherein an endogenous yeast pheromone system receptor protein is not produced in functional form.

10 23. The yeast cell of claim 15, further comprising an indicator gene that produces a detectable signal upon functional coupling of the heterologous G protein coupled receptor to the G protein.

15 24. The yeast cell of claim 9, wherein the cell is a *Saccharomyces cerevisiae* cell.

25. The yeast cell of claim 15, wherein said heterologous G protein coupled receptor is an orphan receptor.

20 26. The yeast cell of claim 4, further comprising an indicator gene that produces a detectable signal upon functional coupling of the heterologous G protein coupled receptor to the G protein.

27. An assay to identify compounds capable of modulating the dissociation of G $\alpha$  and G $\beta\gamma$ , comprising the steps of:

(i) providing a yeast cell according to claim 15 or 26,

25

(ii) contacting the yeast with a test compound; and

(iii) identifying compounds which induce a change in a detectable signal in the yeast cell, wherein said detectable signal indicates dissociation of G $\alpha$  and G $\beta\gamma$ .

28. The assay of claim 27, wherein said test compound is from a library of  
5 non-peptidic organic molecules.

29. A method for identifying a compound which modulates a heterologous G protein coupled receptor, comprising:

10 (i) providing a first, second, third, and fourth yeast cell, each cell comprising:

(a) a G protein, wherein:

15 1) the first yeast cell comprises a first chimeric G protein subunit comprising a first polypeptide from a yeast G protein subunit and a second polypeptide from a mammalian G protein subunit;

20 2) the second yeast cell comprises a second chimeric G protein subunit comprising a first polypeptide derived from a yeast G protein subunit and a second polypeptide from a mammalian G protein subunit, said second chimeric G protein subunit being different from said first chimeric G protein subunit;

25 3) the third yeast cell comprises a third chimeric G protein subunit comprising a first polypeptide from a yeast G protein subunit and a second polypeptide from a mammalian G protein subunit, said third chimeric G protein subunit being different from said first and second chimeric G protein subunits; and

4) the fourth yeast cell comprises an endogenous yeast Gpa1 G protein subunit;

(b) an expressible gene construct encoding a heterologous G protein coupled receptor (GPCR) which couples to the yeast pheromone response pathway; and

5 (c) an indicator gene that produces a change in a detectable signal upon functional coupling of the heterologous GPCR with the G protein;

10 (ii) contacting the first, second, third, and fourth yeast cells with a test compound; and

15 (iii) determining whether the test compound induces a change in a detectable signal in at least one of the first, second, third, or fourth yeast cells to thereby identify a compound which modulates a heterologous GPCR.

30. The assay of claim 29, wherein at least one of said first, second, or third yeast cells comprises a fourth chimeric G protein subunit, said fourth chimeric G protein subunit being different from the first, second, or third chimeric G protein subunit expressed by the first, second, or third yeast cell, respectively.

31. The assay of claim 29, wherein at least one of said first, second, or third chimeric G proteins comprises a first polypeptide from yeast G<sub>p</sub>α1 and a second polypeptide from a mammalian G protein α-subunit.

32. The assay of claim 31, wherein said second polypeptide is from a mammalian G<sub>α</sub>i subunit.

25

33. The assay of claim 31, wherein said second polypeptide is from a mammalian G<sub>α</sub>16 subunit.

34. The assay of claim 30, wherein said second polypeptide is from a mammalian G $\alpha$ s subunit.

5 35. The assay of claim 29, wherein the first chimeric G protein subunit comprises a polypeptide from mammalian G $\alpha$ 12, the second chimeric G protein subunit comprises a polypeptide from mammalian G $\alpha$ 16, and the third chimeric G protein subunit comprises a polypeptide from mammalian G $\alpha$ s.

10 36. The assay of claim 35, wherein the second chimeric G protein subunit comprises G $\alpha$ 16(S270P) and the third chimeric G protein subunit comprises G $\alpha$ s(D229S).

15 37. The assay of claim 29, wherein each of said first, second, and third yeast cells further comprises a fourth chimeric G protein subunit, said fourth chimeric G protein subunit comprising a first polypeptide from yeast STE 18 and a second polypeptide from a mammalian G protein  $\gamma$  subunit.

20 38. The assay of claim 29, wherein the first, second, third, and fourth yeast cells are contacted with each member of a library of test compounds.

39. The assay of claim 37, wherein each member of said library is a non-peptidic organic molecule.

25 40. The assay of claim 29, wherein said first, second, third, and fourth yeast cells are *Saccharomyces cerevisiae* cells.

41. The assay of claim 29, wherein the indicator gene that gives rise to a detectable signal is selected from the group consisting of:  $\beta$  galactosidase, alkaline phosphatase, horseradish peroxidase, exoglucanase, luciferase, BAR1, PHO5, green fluorescent protein, and chloramphenicol acetyl transferase.

5

42. The assay of claim 29, wherein the indicator gene that gives rise to a detectable signal is selected from the group consisting of: HIS 3,  $\beta$  galactosidase, and green fluorescent protein.

10

43. The assay of claim 29, wherein said heterologous G protein coupled receptor is an orphan receptor.

15

44. A recombinant yeast cell which comprises:

a heterologous G protein-coupled receptor (GPCR) expressed in the cell

20

membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast pheromone receptor in a pheromone response pathway of the yeast cell; and

a chimeric G protein subunit comprising an endogenous yeast Gpa1

25

subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell; and wherein modulation of the signal transduction activity of said heterologous GPCR by an extracellular signal provides a detectable signal.

45. A recombinant yeast cell which comprises:

a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast

5 pheromone receptor in a pheromone response pathway of the yeast cell;

a chimeric G protein subunit comprising an endogenous yeast Gpa1 subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous

10 GPCR into the pheromone response pathway of said yeast cell; and

a heterologous polypeptide, wherein the heterologous polypeptide is transported to a location allowing interaction with the extracellular region of the receptor expressed in the cell membrane; and wherein the heterologous polypeptide is expressed at a sufficient level such that modulation of the signal transduction activity of the

15 receptor by the heterologous polypeptide provides a detectable signal.

46. A recombinant yeast cell which comprises:

a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast

20 pheromone receptor in a pheromone response pathway of the yeast cell;

a chimeric G protein subunit comprising an endogenous yeast Gpa1 subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous

25 GPCR into the pheromone response pathway of said yeast cell;

a heterologous polypeptide, wherein the heterologous polypeptide is transported to a location allowing interaction with the extracellular region of the receptor expressed in the cell membrane; and

- a reporter construct that is activated by the pheromone response pathway;
- 5    wherein the heterologous polypeptide is expressed at a sufficient level such that modulation of the signal transduction activity of the receptor by the heterologous polypeptide provides a detectable signal mediated by the reporter construct.

47.    A recombinant yeast cell which comprises:

10      a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast pheromone receptor in a pheromone response pathway of the yeast cell;

15      a chimeric G protein subunit comprising an endogenous yeast G<sub>p</sub>α1 subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell; and

20      a mutation in at least one gene selected from the group consisting of *FARI*, *SST2*, *BARI*, *SVG1*, *STE2*, *STE3*, *STE14*, *MFα1*, *MFα2*, *MFα1* and *MFα2*; and wherein modulation of the signal transduction activity of said heterologous GPCR by an extracellular signal provides a detectable signal.

25      48.    A recombinant yeast cell which comprises:

      a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is

modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast pheromone receptor in a pheromone response pathway of the yeast cell;

- a chimeric G protein subunit comprising an endogenous yeast G<sub>pa1</sub> subunit in which at least the last four C-terminal amino acids are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell;

- a heterologous polypeptide, wherein the heterologous polypeptide is transported to a location allowing interaction with the extracellular region of the receptor expressed in the cell membrane;

a reporter construct that is activated by the pheromone response pathway;

and

- a mutation in at least one gene selected from the group consisting of *FAR1*, *SST2*, *BARI*, *SVGI*, *STE2*, *STE3*, *STE14*, *MFa1*, *MFa2*, *MFa1* and *MFa2*; wherein the heterologous polypeptide is expressed at a sufficient level such that modulation of the signal transduction activity of the receptor by the heterologous polypeptide provides a detectable signal mediated by the reporter construct.

- 20        49.      A mixture of recombinant yeast cells, each cell of which has a cell membrane and each cell of which comprises:

a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal;

a chimeric G protein subunit comprising an endogenous yeast G<sub>pa1</sub> subunit in which at least the last four C-terminal amino acids are replaced with at least

the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell; and

- a heterologous polypeptide, wherein the heterologous polypeptide is  
5 transported to a location allowing interaction with the extracellular region of said receptor expressed in the cell membrane;  
wherein collectively the mixture of cells expresses a library of said heterologous polypeptides, said library being expressible at a sufficient level such that modulation of the signal transduction activity of said receptor by a heterologous polypeptide within the  
10 library provides a detectable signal.

50. A mixture of recombinant yeast cells, each cell of which has a cell membrane and each cell of which comprises:

- a heterologous G protein-coupled receptor (GPCR) expressed in the cell  
15 membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal;

a chimeric G protein subunit comprising an endogenous yeast G<sub>p</sub>α1 subunit in which at least the last four C-terminal amino acids are replaced with at least  
20 the last four C-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell;

a heterologous polypeptide, wherein the heterologous polypeptide is transported to a location allowing interaction with the extracellular region of said  
25 receptor expressed in the cell membrane;

a reporter construct that is activated by the pheromone response pathway;  
and

a mutation in at least one gene selected from the group consisting of *FAR1*, *SST2*, *BARI*, *SVGI*, *STE2*, *STE3*, *STE14*, *MFα1*, *MFα2*, *MFα1* and *MFα2*; wherein the heterologous polypeptide is expressed at a sufficient level such that modulation of the signal transduction activity of the receptor by the heterologous 5 polypeptide provides a detectable signal mediated by the reporter construct.

51. A recombinant yeast cell which comprises:  
a heterologous G protein-coupled receptor (GPCR) expressed in the cell  
membrane of said yeast cell such that signal transduction activity via said receptor is  
modulated by interaction of an extracellular region of the receptor with an extracellular  
signal, said heterologous GPCR acting as a surrogate for an endogenous yeast  
pheromone receptor in a pheromone response pathway of the yeast cell; and  
a chimeric G protein subunit comprising an endogenous yeast G<sub>p</sub>α1  
subunit in which at least the last four C-terminal amino acids of said G<sub>p</sub>α1 are replaced  
with at least the last four C-terminal amino acids of a first heterologous G protein  
subunit, and in which the N-terminus of said G<sub>p</sub>α1 is operably linked to at least the first  
five N-terminal amino acids of a second heterologous G protein subunit, wherein said  
first and second heterologous G protein subunits are the same or different, such that  
expression of said chimeric G protein subunit functionally integrates said heterologous  
20 GPCR into the pheromone response pathway of said yeast cell; and  
wherein modulation of the signal transduction activity of said heterologous GPCR by an  
extracellular signal provides a detectable signal.

52. A mixture of recombinant yeast cells, each cell of which has a cell  
membrane and each cell of which comprises:  
a heterologous G protein-coupled receptor (GPCR) expressed in the cell  
membrane of said yeast cell such that signal transduction activity via said receptor is

modulated by interaction of an extracellular region of the receptor with an extracellular signal;

a chimeric G protein subunit comprising an endogenous yeast G<sub>p</sub>α1 subunit in which at least the last four C-terminal amino acids of said G<sub>p</sub>α1 are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit, and in which the N-terminus of said G<sub>p</sub>α1 is operably linked to at least the first five N-terminal amino acids of a heterologous G protein subunit, such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast cell; and

10 a heterologous polypeptide, wherein the heterologous polypeptide is transported to a location allowing interaction with the extracellular region of said receptor expressed in the cell membrane; wherein collectively the mixture of cells expresses a library of said heterologous polypeptides, said library being expressible at a sufficient level such that modulation of 15 the signal transduction activity of said receptor by a heterologous polypeptide within the library provides a detectable signal.

53. A recombinant yeast cell which comprises:

20 a heterologous G protein-coupled receptor (GPCR) expressed in the cell membrane of said yeast cell such that signal transduction activity via said receptor is modulated by interaction of an extracellular region of the receptor with an extracellular signal, said heterologous GPCR acting as a surrogate for an endogenous yeast pheromone receptor in a pheromone response pathway of the yeast cell; and

25 a chimeric G protein subunit comprising an endogenous yeast G<sub>p</sub>α1 subunit in which at least the last four C-terminal amino acids of said G<sub>p</sub>α1 are replaced with at least the last four C-terminal amino acids of a first heterologous G protein subunit, and in which at least the first five N-terminal amino acids of said G<sub>p</sub>α1 are

replaced with at least the first five N-terminal amino acids of a second heterologous G protein subunit, wherein said first and second heterologous G protein subunits are the same or different; such that expression of said chimeric G protein subunit functionally integrates said heterologous GPCR into the pheromone response pathway of said yeast

5 cell; and

wherein modulation of the signal transduction activity of said heterologous GPCR by an extracellular signal provides a detectable signal.

54. The yeast cell of claim 53, wherein said chimeric G protein subunit  
10 comprises an endogenous yeast G<sub>pα1</sub> subunit in which the last five C-terminal amino acids of said G<sub>pα1</sub> are replaced with the last five C-terminal amino acids of a first heterologous G protein subunit, and in which the first five N-terminal amino acids of said G<sub>pα1</sub> are replaced with the first 11 N-terminal amino acids of a second heterologous G protein  
15 subunits are the same.

55. The yeast cell of claim 54, wherein said chimeric G protein subunit  
comprises an endogenous yeast G<sub>pα1</sub> subunit in which the last five C-terminal amino acids of said G<sub>pα1</sub> are replaced with the last five C-terminal amino acids of a first  
20 heterologous G protein subunit, and in which the first 22 N-terminal amino acids of said G<sub>pα1</sub> are replaced with the first 21 N-terminal amino acids of a second heterologous G protein subunit, wherein said first and second heterologous G protein subunits are the same.

25 56. A chimeric G-protein subunit which comprises an endogenous G<sub>pα1</sub> subunit in which at least the last four C-terminal amino acids of said G<sub>pα1</sub> are replaced with at least the last four C-terminal amino acids of a heterologous G protein subunit.

57. The chimeric G-protein subunit of claim 56, wherein the last five C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last five C-terminal amino acids of a heterologous G protein subunit.

5 58. The chimeric G-protein subunit of claim 57, wherein the last six C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last six C-terminal amino acids of a heterologous G protein subunit.

10 59. A chimeric G-protein subunit which comprises an endogenous G<sub>p</sub>a1 subunit in which at least the last four C-terminal amino acids of said G<sub>p</sub>a1 are replaced with at least the last four C-terminal amino acids of a first heterologous G protein subunit, and in which the N-terminus of said G<sub>p</sub>a1 is operably linked to at least the first five N-terminal amino acids of a second heterologous G protein subunit, wherein said first and second heterologous G protein subunits are the same or different.

15

20 60. The chimeric G-protein subunit of claim 59, in which the last five C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last five C-terminal amino acids of said first heterologous G-protein subunit, and in which the first five N-terminal amino acids of said G<sub>p</sub>a1 are replaced with the first 11 N-terminal amino acids of said second heterologous G protein subunit.

25 61. The chimeric G-protein subunit of claim 59, in which the last five C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last five C-terminal amino acids of said first heterologous G-protein subunit, and in which the first 22 N-terminal amino acids of said G<sub>p</sub>a1 are replaced with the first 21 N-terminal amino acids of said second heterologous G protein subunit.

62. The chimeric G-protein subunit of claim 59, 60 or 61, wherein said first and second heterologous G protein subunits are the same.

5        63. The yeast cell of claim 44, 48, 51, 54 or 55, wherein said heterologous G protein-coupled receptor is selected from the group consisting of human Melatonin receptor 1a, human Melatonin receptor 1b, human Nociceptin receptor, human Somatostatin receptor 2, human Formyl peptide receptor 1, human Galanin receptor 1, Rabbit CXCR1, human C5a receptor, human Adenosine receptor 2a, Rat vasoactive 10 intestinal receptor, human Melanocortin receptor 4, Rat Neurotensin receptor, human Formyl peptide like receptor 1, human CXCR2 (mutant), human CXCR4, human Adenosine 2b receptor, human bombesin receptor subtype 3 (BRS3), human Bradykinin receptor 2, and orphan receptors.

15        64. The yeast cell of claim 44, 48, 51, 54 or 55, wherein said heterologous G protein subunit is selected from the group consisting of  $G\alpha_s$ ,  $G\alpha_i2$ ,  $G\alpha_i3$ ,  $G\alpha_q$ ,  $G\alpha_o$ ,  $G\alpha_{o_b}$  and  $G\alpha_{16}$ .

20        65. The yeast cell of claim 44, 48, 51, 54 or 55, wherein said heterologous G protein subunit is non-mammalian.

25        66. The yeast cell of claim 2, 44, 45, 46, 47, 48, 49 or 50, wherein from the last 4 to the last 40 C-terminal amino acids of said  $G\beta\gamma 1$  are replaced with from the last 4 to the last 40 C-terminal amino acids of said heterologous G-protein unit.

67. The yeast cell of claim 66, wherein from the last 5 to the last 6 C-terminal amino acids of said G<sub>p</sub>a1 are replaced with from the last 5 to the last 6 C-terminal amino acids of said heterologous G-protein unit.

5 68. The yeast cell of claim 67, wherein the last 5 C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last 5 C-terminal amino acids of said heterologous G-protein unit.

10 69. The yeast cell of claim 67, wherein the last 6 C-terminal amino acids of said G<sub>p</sub>a1 are replaced with the last 6 C-terminal amino acids of said heterologous G-protein unit.

15 70. The yeast cell of claim 51, wherein from the last 4 to the last 40 C-terminal amino acids of said G<sub>p</sub>a1 are replaced with from the last 4 to the last 40 terminal amino acids of said first heterologous G protein subunit, and wherein the N-terminus of said G<sub>p</sub>a1 is operably linked to from the first 5 to the first 40 N-terminal amino acids of said second heterologous G protein subunit.

20 71. The yeast cell of claim 70, wherein from the last 5 to the last 6 C-terminal amino acids of said G<sub>p</sub>a1 are replaced with from the last 5 to the last 6 C-terminal amino acids of said first heterologous G-protein subunit, and wherein from the first 1 to first 50 N-terminal amino acids of said G<sub>p</sub>a1 are replaced with from the first 6 to the first 40 N-terminal amino acids of said second heterologous G protein subunit.

25 72. The yeast cell of claim 71, wherein the first 5 N-terminal amino acids of said G<sub>p</sub>a1 are replaced with the first 11 N-terminal amino acids of said second heterologous G protein subunit.

73. The yeast cell of claim 71, wherein the first 22 N-terminal amino acids of said G<sub>p</sub>a1 are replaced with the first 21 N-terminal amino acids of said second heterologous G protein subunit.

5 74. The yeast cell of claim 70, 71, 72 or 73, wherein said first and second heterologous G protein subunits are the same.

75. The yeast cell of claim 74, wherein said heterologous G protein subunits are mammalian.

10

76. The yeast cell of claim 75, wherein said heterologous G protein subunits are G<sub>a</sub>q subunits.

15 77. The yeast cell of claim 1, 2, 3, 4, 44, 47, 51 or 53, said cell further comprising a heterologous polypeptide, wherein the heterologous polypeptide is transported to a location allowing interaction with the extracellular region of the receptor expressed in the cell membrane; and wherein the heterologous polypeptide is expressed at a sufficient level such that modulation of the signal transduction activity of the receptor by the heterologous polypeptide provides a detectable signal.

20

78. The yeast cell of claim 45, 46, 48, 49, 50, 52 or 77, wherein said heterologous polypeptide is a ligand of said heterologous GPCR.

25 79. The yeast cell of claim 78, said cell further comprising a second heterologous polypeptide, said second heterologous polypeptide also being transported

to a location allowing interaction with the extracellular region of said heterologous GPCR expressed in the cell membrane.

80. The yeast cell of claim 78, wherein the heterologous polypeptide includes  
5 a signal sequence that facilitates transport of the polypeptide to a location allowing interaction with the extracellular region of the receptor.

81. The yeast cell of claim 80, wherein the signal sequence corresponds to a leader peptide of the *Saccharomyces cerevisiae*  $\alpha$  factor or  $\alpha$ -factor.

10

82. The yeast cell of claim 1, 2, 3, 44, 45, 46, 49, 51, 52 or 53, further comprising a mutation in at least one gene selected from the group consisting of *FAR1*, *SST2*, *BARI*, *SVG1*, *STE2*, *STE3*, *STE14*, *MFa1*, *MFa2*, *MFa1* and *MFa2*.

15

83. The yeast cell of claim 46, 47, 49 or 82, wherein said yeast cell has a *STE14* mutation.

84. The yeast cell of claim 83, wherein said cell has a *STE2* or *STE3* mutation.

20

85. The yeast cell of claim 1, 2, 3, 4, 9, 43, 46, 50 or 52, said cell further comprising a reporter construct that is activated by the pheromone response pathway, thereby providing the detectable signal.

25

86. The yeast cell of claim 44, 48 or 51, said cell further comprising a reporter construct that is activated by the pheromone response pathway, wherein the heterologous polypeptide is expressed at a sufficient level such that modulation of the

signal transduction activity of the receptor by the heterologous polypeptide provides a detectable signal mediated by the reporter construct.

87. The yeast cell of claim 46, 48, 50, 86 or 87, wherein the reporter  
5 construct comprises a pheromone-responsive promoter operably linked to a reporter  
gene.

88. The yeast cell of claim 87, wherein said reporter gene is selected from the  
group consisting of  $\beta$ -galactosidase (LacZ), alkaline phosphatase, horseradish  
10 peroxidase, exoglucanase, luciferase, BAR1, PHO5, HIS3, green fluorescent protein  
(GFP), and chloramphenicol acetyl transferase.

89. The yeast cell of claim 88, wherein the pheromone-responsive promoter  
is the FUS1 promoter.  
15

90. The yeast cell of claim 89, wherein said FUS1 promoter is operably  
linked to a reporter gene selected from the group consisting of LacZ and GFP.

91. The yeast cell claim 90, wherein said FUS1 promoter is operatively  
20 linked to GFP.

92. The yeast cell of claim 91, wherein said detectable signal is generated by  
production of GFP.

25 93. The yeast cell of claim 1, 2, 3, 4, 44, 45, 46, 47, 48, 49, 50, 51, 52 or 53,  
which belongs to the species *Saccharomyces cerevisiae*.

94. A method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

contacting a mixture of yeast cells as claimed in claim 1, 2, 3, 44, 47, 51, or 53 with a test compound;

5 allowing cells within the mixture to generate a detectable signal; and identifying the test compound as a modulator of said receptor.

95. A method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

10 contacting a mixture of yeast cells as claimed in claim 45, 46, 47, 48, 49 or 52 with a ligand of said receptor;

allowing cells within the mixture to generate a detectable signal; and identifying the heterologous polypeptide expressed by the yeast cells as a modulator of said receptor.

15

96. A method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

providing a mixture of yeast cells as claimed in claim 79;  
allowing cells within the mixture to generate a detectable signal; and  
20 identifying the second heterologous polypeptide as modulator of said receptor.

97. A method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

25 contacting a first mixture of yeast cells as claimed in claim 78 with a second mixture of yeast cells, wherein collectively the second mixture of yeast cells expresses a library of heterologous test polypeptides that are transported to a location

allowing interaction with the extracellular region of said receptor expressed in the cell membrane of the yeast cells of the first mixture;

allowing cells within the first mixture to generate a detectable signal; and  
identifying a heterologous test polypeptide within the second mixture as a  
5 modulator of said receptor expressed by the yeast cell.

98. A method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

providing a mixture of recombinant yeast cells as claimed in claim 46,  
10 47, 48, 49, 50, 51 or 77 wherein the heterologous polypeptide in each of said yeast cells is a heterologous test polypeptide;  
allowing cells within the mixture to generate a detectable signal; and  
identifying the heterologous test polypeptide as a modulator of said heterologous receptor.

15

99. A method for identifying a modulator of a heterologous G protein-coupled receptor expressed by a yeast cell, comprising:

contacting a mixture of yeast cells as claimed in claim 45, 46, 48, 49, 50,  
20 52 or 77 with a ligand of said receptor;  
allowing cells within the mixture to generate a detectable signal; and  
identifying the heterologous polypeptide expressed by the yeast cells as a modulator of said receptor.

100. The method of claim 93, wherein each yeast cell of said mixture further  
25 comprises a reporter construct that is activated by the pheromone response pathway,  
thereby providing the detectable signal.

101. The method of claim 99, wherein the reporter construct comprises a pheromone-responsive promoter operably linked to a reporter gene.

102. The method of claim 100, wherein the pheromone-responsive promoter is  
5 the FUS1 promoter.

103. The method of claim 101, wherein said FUS1 promoter is operably linked to a reporter gene selected from the group consisting of LacZ and GFP.

10 104. The method of claim 103, wherein said FUS1 promoter is operatively linked to GFP.

15 105. The method of claim 104, wherein said detectable signal is generated by producer of GFP.

106. The method of claim 94, 95, 96, or 97, wherein the heterologous polypeptide includes a signal sequence that facilitates transport of the polypeptide to a location allowing interaction with the extracellular region of the receptor.

20 107. The method of claim 105, wherein the signal sequence corresponds to a leader peptide of the *Saccharomyces cerevisiae*  $\alpha$  factor or  $\alpha$ -factor.

108. The method of claim 93, 94, 95, 96, 97 or 98, wherein said yeast cells belong to the species *Saccharomyces cerevisiae*.

25 109. The chimeric G protein subunit comprising the amino acid sequence of SEQ ID NO: 107.

110. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 108.

5 111. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 109.

112. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 110.

10

113. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 111.

15

114. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 112.

115. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 113.

20

116. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 114.

117. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 115.

25

118. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 118.

**119. The chimeric G protein subunit comprising the amino acid sequence of  
SEQ ID NO: 123.**

# FIG. 1

Sequence alignments of N-terminal regions of Ga subunits and N-terminal sequences of GPA<sub>41</sub>- Ga hybrid proteins.

## A. Alignment of GPA1 with Ga subunits

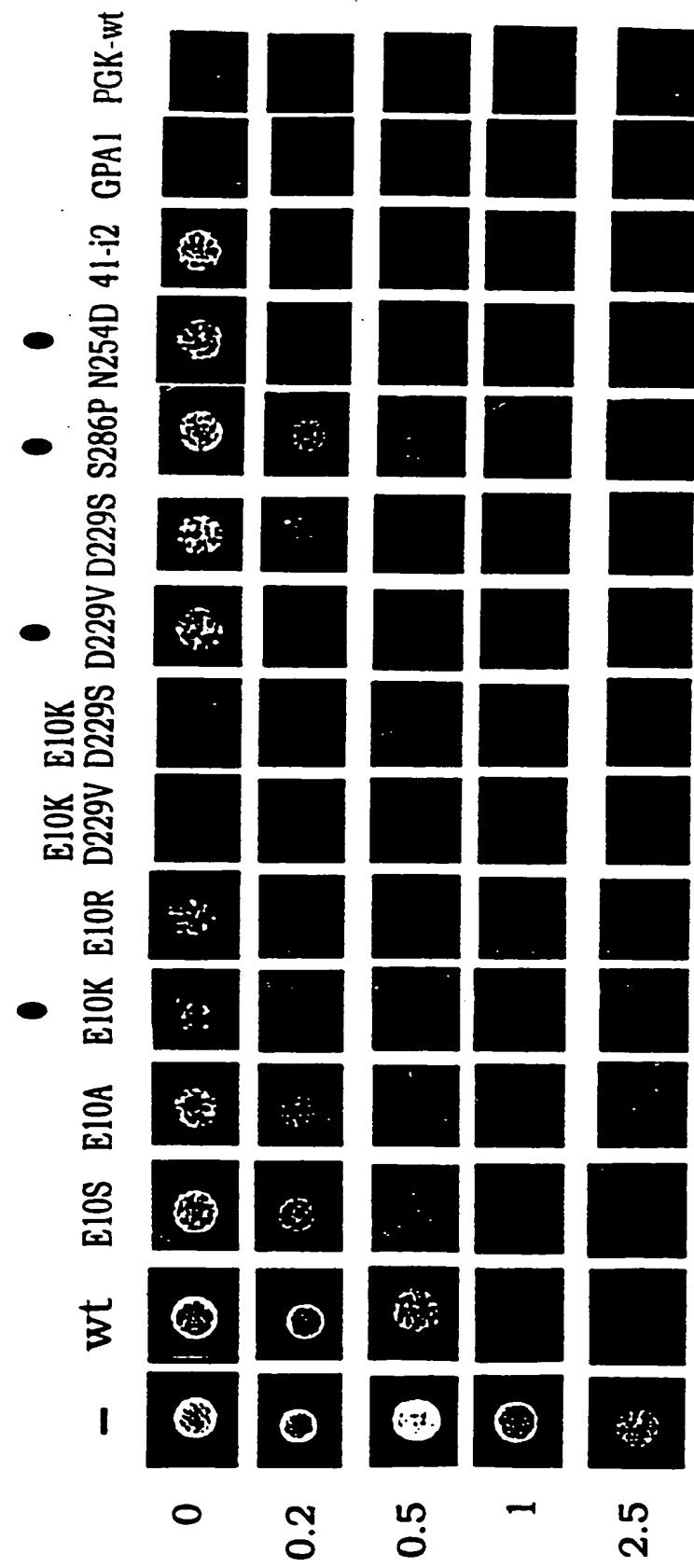
GPA1	MGC . TVSTQTIGDESDPFLQNKRANDVIEQSLQLEKORDKNEIKLLLGGAGESGKSTIVKQLKLLHQ. . . .
GaS	MGC LCTS . KTEDQRNEEKAQREANKKIEQLQDKQVYRATHRLLLGGAGESGKSTIVKQMRILHV. . . .
Gai2	MGC . TVS . . . . AEDKAAAERSKMDKNLREDGEKAAREVKLLLGGAGESGKSTIVKQMKIIHE. . . .
Gai3	MGC . TVS . . . . AEDKAVERS SKMIDRNLRDGEKAKEVAKL LGGAGESGKSTIVKQMRKIIHE. . . .
Ga16	MARSILTWRCCCPWCCLTEDEKAARVDQEINRILLEQQKQDRGELKLLLGPGESGKSTIFIKQMRITHG. . . .

## B. GPA<sub>41</sub>- Ga Hybrids

GPA <sub>41</sub> - GaS	MGC . TVSTQTIGDESDPFLQNKRANDVIEQSLQLEKORDKNERKL LGGAGESGKSTIVKQMRILHV. . . .
GPA <sub>41</sub> - Gai2	MGC . TVSTQTIGDESDPFLQNKRANDVIEQSLQLEKORDKNEVKLLLGGAGESGKSTIVKQMRKIIHE. . . .
GPA <sub>41</sub> - Gai3	MGC . TVSTQTIGDESDPFLQNKRANDVIEQSLQLEKORDKNEVKLLLGGAGESGKSTIVKQMRKIIHE. . . .
GPA <sub>41</sub> - Ga16	MGC . TVSTQTIGDESDPFLQNKRANDVIEQSLQLEKORDKNEKL LLLGPGESGKSTIFIKQMRITHG. . . .

2 / 2

FIG. 2



Coupling of Gas mutants to yeast  $\text{G}\beta\gamma$   
mM AT ↪

# INTERNATIONAL SEARCH REPORT

Int. Application No  
PCT/US 98/21168

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC 6	C12N15/12	C12N15/31	C12N15/62	C12N1/19	C07K14/47
	C07K14/395	C1201/02	C1201/68	G01N33/50	

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C12N C07K C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>Y.-S. KANG ET AL.: "Effects of expression of mammalian Galpha and hybrid mammalian yeast Galpha proteins on the yeast pheromones response signal transduction pathway"            J. MOL. BIOL.,            vol. 10, no. 6, June 1990, pages            2582-2590, XP002092715            ACADEMIC PRESS, US            cited in the application            see the whole document</p> <p>---</p>	1-119
A	<p>WO 92 05244 A (UNIV DUKE) 2 April 1992            cited in the application            see the whole document</p> <p>---</p> <p>---</p>	1-119



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

9 February 1999

Date of mailing of the international search report

18/02/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.  
 Fax: (+31-70) 340-3016

Authorized officer

Hornig, H

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 98/21168

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 94 23025 A (CADUS PHARMACEUTICALS INC) 13 October 1994 see the whole document ----	1-119
A	WO 97 11159 A (HEARTLAND BIOTECHNOLOGIES L L ;DAS (US); MANDELL ROBERT BARRY (US)) 27 March 1997 see the whole document ----	1-119
A	WO 95 30012 A (CADUS PHARMACEUTICAL CORP ;BROACH JAMES R (US); MANFREDI JOHN P (U) 9 November 1995 see the whole document ----	1-119
P,A	WO 98 13513 A (CADUS PHARMACEUTICAL CORP) 2 April 1998 see the whole document ----	1-119

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/21168

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9205244	A 02-04-1992	AU 652576	B	01-09-1994
		AU 8511591	A	15-04-1992
		CA 2092717	A	14-03-1992
		EP 0548165	A	30-06-1993
		JP 6500693	T	27-01-1994
		US 5482835	A	09-01-1996
		US 5739029	A	14-04-1998
-----				
WO 9423025	A 13-10-1994	AU 6354198	A	09-07-1998
		AU 685103	B	15-01-1998
		AU 6490994	A	24-10-1994
		CA 2158274	A	13-10-1994
		EP 0692025	A	17-01-1996
		JP 8510115	T	29-10-1996
		US 5789184	A	04-08-1998
-----				
WO 9711159	A 27-03-1997	AU 7115896	A	09-04-1997
		CA 2232733	A	27-03-1997
		EP 0862614	A	09-09-1998
-----				
WO 9530012	A 09-11-1995	AU 2635495	A	29-11-1995
		EP 0758392	A	19-02-1997
		JP 9512437	T	16-12-1997
-----				
WO 9813513	A 02-04-1998	AU 4593097	A	17-04-1998